2011

Self-Efficacy in Music Performance: Measuring the Sources Among Secondary School Music Students

Michael S. Zelenak
University of South Florida, zekesbch@tampabay.rr.com

Self-Efficacy in Music Performance:

Measuring the Sources among Secondary School Music Students

by

Michael S. Zelenak

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
School of Music
College of The Arts
University of South Florida

Major Professor: C. Victor Fung, Ph.D.
David A. Williams, Ph.D.
Janet Moore, Ed.D.
Robert F. Dedrick, Ph.D.
Clint Randles, Ph.D.

Date of Approval:
July 13, 2011

Keywords: music education, self-esteem, music aptitude, validation,
sources of self-efficacy

Dedication

Nothing is accomplished without the love and support of others. In return, I would like to dedicate this work to everyone who has helped me along my life's journey. There is a part of you residing in some corner of this work.

In particular, I would like offer this work in memory of my father Mike whose encouragement to "be the best" continues to provide a source of motivation and to my mother Joann who has always been there with an open ear and a warm heart.

Acknowledgements

I would like to acknowledge the members of the USF music education faculty, Victor Fung, David Williams, Janet Moore, Clint Randles, Jack Heller, former member Carlos Rodriguez, and Bill Hayden for their personal contributions to my professional development. They opened new intellectual doors for me and provided examples of academic excellence. In addition, I would like to recognize my fellow students in the doctoral program who have been a constant source of ideas and who have demonstrated great patience as I formulated my thoughts and practiced my presentations. I would also like to identify Robert Dedrick, Jeffrey Kromrey, John Ferron, and Constance Hines from the USF College of Education's Research and Measurement Department for captivating my imagination with statistical analysis.

I cannot forget the music teachers who participated in this project. Although your names must remain anonymous, your willingness to participate indicated a true dedication to your students and the field of music education. I would also like to acknowledge the school districts that were open to new ideas about music education and allowed me into their classrooms.

Finally, I would like to recognize the members of my family who provided me with a shoulder to lean on when I was down, and a swift kick when I needed to get back up. Special thanks to Cory and Michael; Julie and Dave; John and Veronica; as well as, Carrie, Jerry, Jamie, Mariah, and the angel among us, Alena.

You brought out the best in me.

Table of Contents

List of Tables iii

List of Figures v

Abstract vi

Chapter 1: Introduction 1
- Theoretical Framework 5
- Purpose 10
- Research Questions 11
- Significance of the Study 12
- Operational Definition of Terms 12
  - Independent variables 12
  - Dependent variables 13
- Delimitations 14
- Limitations 15

Chapter 2: Review of the Literature 17
- Self-Efficacy 17
- Self-Efficacy Scales 22
  - Guidelines for constructing self-efficacy scales 23
  - General scales of self-efficacy 23
  - Self-efficacy scales in education 25
  - Sources of self-efficacy scales 27
  - Validity 36
  - Reliability 38
  - Response Formats 38
  - Age and grade differences 42
- Self-Efficacy in Music Education 44
- Summary 48

Chapter 3: Method 50
- Research Design 50
- Population and Sample 52
- Instruments 57
- Data Collection 68
- Data Analysis 70
  - Software 79

Chapter 4: Results 80
Data Characteristics 80
Missing data 80
Outliers 81
Reliability 82
Item Evaluation of the Music Performance Self-Efficacy Scale 84
Analytical Assumptions 91
Research Questions 95
Question 1 95
Question 2 98
Question 3 99
Question 4 103
Question 5 103
Question 6 107
Question 7 109
Summary 113

Chapter 5: Summary, Discussion, and Recommendations 116
Summary 116
Discussion 117
Fundamental knowledge 117
Validity and reliability of the Music Performance Self-Efficacy Scale 122
Recommendations 123
For Further Research 124
For Music Educators 126

List of References 131

Appendices 141
Appendix A: Demographic Information 142
Appendix B: Music Performance Self-Efficacy Scale 143
Appendix C: Music Performance Self-Efficacy Scale (Preliminary Study) 145
Appendix D: Sources of Mathematics Self-Efficacy Scale 148
Appendix E: Self-Esteem of Musical Ability 150
Appendix F: Advanced Measures of Music Audiation 153
Appendix G: Self-Efficacy Rating Form 156
Appendix H: Instruction for Self-Efficacy in Music Performance Study 157
Appendix I: Introduction Script 159
Appendix J: Parental Permission to Participate in Research 161
Appendix K: Assent to Participate in Research 165
Appendix L: Informed Consent to Participate in Research 168
Appendix M: Using Excel to Calculate Self-Efficacy Scores 171
Appendix N: Inter-Item Correlation Coefficients 173
Appendix O: Means, Standard Deviations, and Cell Sizes for MPSES 175

About the Author ..... End Page

List of Tables

Table 1 General Descriptions of Sources of Self-Efficacy Scales 28

Table 2 Internal Consistencies (Cronbach's α) of Sources of Self-Efficacy Scales 30

Table 3 Correlations of Sources of Self-Efficacy Scores with Subject Self-Efficacy Scores 32

Table 4 Comparison of Standardized Regression Coefficients among Self-Efficacy Scales 34

Table 5 Response Rates by School 54

Table 6 Demographic Characteristics of Participants 56

Table 7 Correlations of Sources of Self-Efficacy Scores and Composite Self-Efficacy Scores from Preliminary Study 60

Table 8 Internal Consistency Estimates 84

Table 9 MPSES Item Statistics Grouped by Source of Self-Efficacy 85

Table 10 Descriptive Statistics for Sources of Self-Efficacy 88

Table 11 Correlations between the Sources of Self-Efficacy and Composite Self-Efficacy 91

Table 12 Data Distribution Indicators for All Scales 93

Table 13 Estimated Standardized Loadings of Sources of Self-Efficacy onto Items 98

Table 14 Testing for Factor Invariance among 24 MPSES Items across Middle (n = 150) and High School (n = 140) Music Students 102

Table 15 MANOVA Results from Music Performance Self-Efficacy Scale 104

Table 16 Univariate F-test for Mastery Experience 105

Table 17 Univariate *F*-test for Vicarious Experience 106

Table 18 Univariate *F*-test for Verbal/Social Persuasion 106

Table 19 Univariate *F*-test for Physiological State 107

Table 20 Results from the Music Performance Self-Efficacy Scale by Ensemble 127

List of Figures

Figure 1. Self-Efficacy in Bandura's Model of Triadic Reciprocal Causation 6

Figure 2. Bandura's Social Cognitive Theory Illustrated as Theoretical Pathways....7

Figure 3. Model Identification of the Music Performance Self-Efficacy Scale.....72

Figure 4. Standardized Loadings and Standard Errors from Self-Efficacy onto Sources of Self-Efficacy 99

Figure 5. Regression Results Predicting Music Self-Efficacy from Music Aptitude.....109

Abstract

The purpose of this study was to develop a greater understanding of self-efficacy in music performance. I sought to (a) contribute to fundamental knowledge of self-efficacy in music performance, (b) determine whether scores from the Music Performance Self-Efficacy Scale (MPSES) were valid and reliable, and (c) provide insights for developing self-efficacy among secondary school music students. Participants (N = 290) were middle and high school students in band, chorus, and string orchestra ensembles from 10 schools in the southeast and western regions of the United States. Participants completed four online questionnaires: (a) Music Performance Self-Efficacy Scale, (b) Sources of Middle School Mathematics Self-Efficacy Scale, (c) Self-Esteem of Musical Ability, and (d) Advanced Measures of Music Audiation. Teachers provided 5-point Likert-type ratings of their student's music self-efficacy.

Data from the MPSES demonstrated good fit with Bandura's proposed self-efficacy model (SRMR = .06, RMSEA = .06). The strength of the relationships between the sources and composite construct were consistent with theory and findings from other studies. Mastery experience was strongest, followed by verbal/social persuasion, physiological state, and vicarious experience. No differences in responses were found between middle and high school students, or among band, chorus, and string orchestra students. Items on the MPSES were also found to assess participants equally across grade levels.

Music aptitude was found to predict self-efficacy in music performance (β = .16) and accounted for 3% of the variance in self-efficacy. This finding raises questions about the role of self-efficacy in mediating the relationship between music aptitude and music achievement.

The Music Performance Self-Efficacy Scale was found to be a valid and reliable measure of self-efficacy in music performance. Evidence of validity was based on test content, response process, and internal structure, along with convergent, discriminant, and multi-method relationships. Evidence of reliability was based on test-retest correlation (r = .87) and internal consistency (α = .88).

Recommendations for researchers were to link the sources of self-efficacy to performance achievement; explore self-efficacy's relationship with aptitude and achievement; and expand the investigation to other populations. Recommendations for educators focused on improving understanding of self-efficacy and interpreting MPSES results.

Chapter 1: Introduction

Individuals hold self-perceptions of their music performance abilities. These self-perceptions, however, are frequently overlooked by educators in the instruction they deliver to middle and high school students in band, chorus, and string orchestra classes. Instruction in these ensembles typically focuses on the development of students' psychomotor skills related to performing music and ignores the self-perceptions associated with those skills. Students spend hours sharpening their physical coordination (Rohwer & Polk, 2006; Sehman, 2000), aural acuity (Kendall, 1988), and visual discernment (Hayward & Gromko, 2009) while taking little time to nurture their beliefs in those abilities.

Researchers have acknowledged this imbalance. Greenberg (1970) proposed that underachievement in music was a result of a low self-concept of one's ability. More recently, Randles (2011) discovered that as grade level increased, students' self-perception of being a good musician decreased. Additional symptoms have included performance anxiety among student musicians (Abel & Larkin, 1990; Craske & Craig, 1984; Papageorgi, Hallam, & Welch, 2007) and low enrollment in school music programs (Austin & Vispoel, 1998; McPherson & Hendricks, 2010; Williams, 2007). The preoccupation among music educators with developing psychomotor skills and disregarding the self-perceptions associated with those skills has had negative effects on music students.

In contrast, a positive relationship has been found between the self-perception of self-efficacy and achievement in music performance. McPherson and McCormick (2000, 2006) and McCormick and McPherson (2003) reported that performance self-efficacy was a better predictor of achievement in music performance than intrinsic value, general self-efficacy, and self-regulation. In light of this finding, many music educators still fail to address self-efficacy in their instruction. This inattentiveness may reflect a lack of knowledge, the absence of a valid and reliable measure, or inadequate strategies for developing self-efficacy.

Knowledge and measurement, however, are intertwined. Theoretical investigations require the use of valid measurement instruments and valid instruments must be built on unassailable theories. Like the proverbial chicken and egg, one is dependent on the other. Cronbach and Meehl (1955) stated, "Confidence in theory is increased as more relevant evidence confirms it ..." (p. 265). They also advised, "A test should not be used to measure a trait until its proponent establishes that predictions made from such measures are consistent with the best available theory of the trait" (p. 262). With this impasse in mind, I determined the need in this study to address both sides concurrently. I used the Music Performance Self-Efficacy Scale (Zelenak, in press) to investigate Bandura's (1986) social cognitive theory in the domain of music performance. And, I used Bandura's (1986) social cognitive theory to determine the validity and reliability of scores generated by the Music Performance Self-Efficacy Scale. The investigation of one was not possible without an evaluation of the other.

Bandura (1986, 1997) identified self-efficacy as a primary determinant of human behavior in his social cognitive theory. He described self-efficacy as self-referent thought

through which individuals assessed their skills and abilities to accomplish specific tasks. Self-efficacy mediated the relationship between cognition and behavior and affected human functioning on many levels. It influenced actions taken, the level of effort put forth, the amount of time that individuals persevered, the degree of resiliency to recover from setbacks, the direction of thought patterns, the magnitude of stress and anxiety, and the level of achievement. There was one characteristic of self-efficacy in particular that has attracted educators to this construct—it was malleable (Berry & West, 1993).

Bandura identified four sources of information that contributed to the development of self-efficacy beliefs (1977, 1986). Those sources were (a) enactive mastery experience, (b) vicarious experience, (c) verbal/social persuasion, and (d) physiological state. Mastery experience referred to an individual's past experiences of success or failure while participating in the activity, vicarious experience referred to the observations of others similar to the individual engaged in the task, verbal/social persuasion referred to judgments or opinions provided by others, and physiological state referred to the degree and quality of arousal brought on by engagement in the task. The Music Performance Self-Efficacy Scale (Zelenak, in press) was designed to measure these four sources of self-efficacy among secondary school music students (i.e., grades 6-12, or ages 11-18 years) participating in band, chorus, and string orchestra classes.

Along with self-efficacy, music aptitude has been found to influence achievement in music performance (Fullen, 1993; Gordon, 1989). According to Gordon (1993), music aptitude was "a measure of a student's potential to learn music" and music achievement was "a measure of what a student has learned" (p. 2). Fullen (1993) conducted a study among choral music students grades 7-12. He administered Gordon's (1989) Advanced

Measures of Music Audiation as a pre-instructional measure of music aptitude and collected judges' ratings of student vocal performances as a post-instructional measure of music achievement. He calculated the predictive validity coefficient between music aptitude and music achievement to be $r = .25$ among 7th and 8th grade students and $r = .24$ among 9th -12th grade students. These correlations represent the relationship between scores on the AMMA and the judges' performance ratings. The results of Fullen's study along with the findings of McPherson and McCormick provided a rationale for investigating the relationship between music aptitude and self-efficacy in music performance in the current study. I am aware of no other investigation of this relationship. Examining the relationship between music aptitude and self-efficacy in music performance extended the findings of Fullen (1993) and Gordon (1989) to include a mediating variable (i.e., self-efficacy) between music aptitude and music performance achievement. It linked the findings of McPherson and McCormick to the work of Fullen and Gordon.

The Music Performance Self-Efficacy Scale (MPSES) (Zelenak, in press) was constructed to measure the sources of information that contribute to the development of self-efficacy in music performance. To determine the effectiveness of the scale, it was necessary to collect evidence to establish the validity and reliability of the scores produced by the scale. Previously, the MPSES had been evaluated with a small group of students in one school (Zelenak, in press). The scale, however, needed to be administered to a larger and more diverse sample of participants to evaluate its performance with other populations.

The current study followed recommendations made by the American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME) (1999) to assess the validity and reliability of measurement instruments used in educational settings. They defined validity as "the degree to which evidence and theory support the interpretations of test scores entailed by the proposed uses of tests" (p. 9). They proposed that evidence of validity should be based on test content, response processes, internal structure, relations to other variables, and convergent and discriminant sources. In the same publication, these organizations also provided a definition for reliability. They defined reliability as "the consistency of such measurements when the testing procedure is repeated on a population of individuals or groups" (p. 25). Establishing evidence of validity and reliability for scores generated by the MPSES was necessary before the MPSES could be recommended for use in future applications.

Theoretical Framework

Bandura's (1986) social cognitive theory provided a clear-cut framework on which to base this investigation of self-efficacy. In his theory, Bandura proposed a model of reciprocal causation between cognition, behavior, and environment. The reciprocal nature of his model contradicted previous unidirectional models such as behaviorism (Skinner, 1953). In unidirectional models, thought and behavior were driven by positive outcomes, or rewards. In Bandura's model however, thought and behavior were determined by the interaction of cognition, behavior, and environment. Self-efficacy moderated the relationship between cognition and behavior, allowing individuals to cognitively evaluate how well their abilities match the demands of the task (Figure 1).

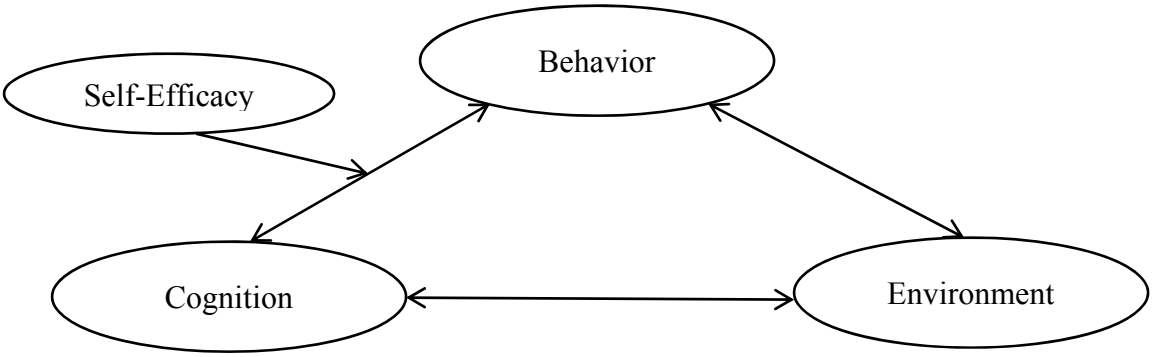

Figure 1. Self-Efficacy in Bandura's Model of Triadic Reciprocal Causation.

Education researchers have incorporated the principles of Bandura's social cognitive theory into the construction of various behavioral models. In one example, Lent, Brown, and Hackett (1994) built a model to explain and predict the development of career intentions (Figure 2). In their model, they illustrated the sources of self-efficacy as having a direct effect on self-efficacy and outcome expectations. Continuing through the model, self-efficacy and outcome expectations then had direct and indirect influences on interests, intentions, activity selection, and performance attainments. Direct influence referred to the relationship between variables in which a change in the value of one variable resulted in a consistent change in the other, while an indirect influence indicated the presence of a moderator variable that altered the relationship between the two variables. In Lent and others' model, self-efficacy had a direct and an indirect influence on intentions. The indirect influence was mediated by the variable interest. The indirect influence of self-efficacy on intentions through the mediating variable made the influence of self-efficacy on intentions different than if it had not been mediated by interest.

Changes in the magnitude of intention did not correspond exactly with changes in the magnitude of self-efficacy because interest altered the relationship between them. Following these direct and indirect influences, the outcome variable performance attainments then had a reciprocal effect on the sources of self-efficacy. Lent and others' model confirmed the reciprocal nature of Bandura's theory.

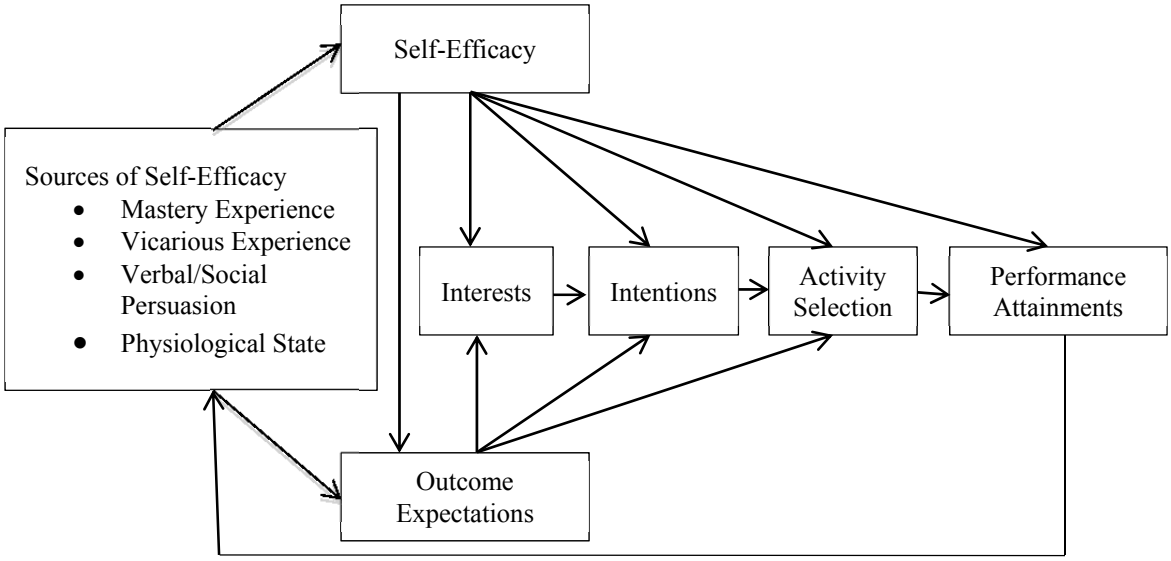

Figure 2. Bandura's Social Cognitive Theory Illustrated as Theoretical Pathways. Adapted from "Toward a Unifying Social Cognitive Theory of Career and Academic Interest, Choice, and Performance" by R. W. Lent, S. D. Brown, and G. Hackett, 1994, *Journal of Vocational Behavior, 45*, 79-122. Copyright 1993 by R. W. Lent, S. D. Brown, and G. Hackett.

Fouad and Smith (1996) tested Lent and others' (1994) model with middle school students and mathematics interest. They examined the fit of data generated by their researcher-designed assessment with Lent and others' model using Byrne's (1989) benchmarks. Although results from the χ^2 statistic suggested a misfit between the data and the model, χ^2 (4, N = 380) = 19.01, $p < .05$, acceptable levels of fit were indicated by

the adjusted goodness of fit index (AGFI = .96) and the root mean square residual (RMR = .02). In confirmatory factor analysis, the χ^2 statistic is sensitive to sample size. As the sample size increases, the likelihood of misfit also increases. As an alternative, the AGFI and RMR values are often used to evaluate fit. Although both values are considered absolute indices, their values represent different properties of the comparison between the sample matrix and the proposed model matrix. The AGFI is a comparison of the overall variance and covariance in the sample matrix that can be accounted for by the proposed model matrix. The closer the AGFI value is to 1, the greater the similarity of the matrices. The RMR is the square root of the mean squared residuals (i.e., errors) between the sample and proposed model matrices. The less difference there is between the matrices, the smaller the residuals will be, resulting in a smaller RMR value (Stevens, 2009). Values above .90 for the AGFI and values below .05 for the RMR indicate the sample matrix is similar to the proposed model matrix.

Further evidence has supported the use of Bandura's theory in constructing models of human behavior. In the same study, Fouad and Smith (1996) conducted a structural equation analysis using Lent and others' (1994) model and found that path coefficients were significant ($p < .05$) for all pathways except self-efficacy to interest and outcome expectation to interest among middle school students. In a similar study, Lent, Lopez, and Bieschke (1991) found significant correlations among self-efficacy, outcome expectations, interests, and mathematics performance with college undergraduates. Although not based in the domain of music, these studies provided evidence supporting the use of Bandura's theory as a framework for understanding self-efficacy and human behavior.

Music education researchers have investigated self-efficacy in music performance. McPherson and McCormick (2000) asked music students ($N = 349$) 9-18 years of age to complete a researcher-designed self-report questionnaire before participating in a music performance examination. They found that performance self-efficacy accounted for the largest percentage of variance in each of three age groups: Group 1 ($M = 11.62$ years) – 18%, Group 2 ($M = 13.4$ years) – 28%, and Group 3 ($M =$ 15.11 years – 23%). This finding led the authors to conclude that performance self-efficacy had a greater influence on music performance achievement than did intrinsic value, general self-efficacy, or self-regulation. In the process of stepwise regression analysis, performance self-efficacy was entered as the first step in the equation followed by the other variables. The authors suggested that this finding "highlights the importance for students to enter a music examination with a positive belief in their own capacity to succeed" (p. 37). McCormick and McPherson (2003) replicated the previous study using a different group of participants and obtained similar results.

Along with regression, structural equation modeling has been used to examine the relationships between self-efficacy and related constructs. McCormick and McPherson (2003) constructed several models to investigate the relationships among anxiety, self-regulation, grade level, practice time, cognitive strategy use, formal practice, informal practice, self-efficacy, and performance. The data, however, did not fit any of the proposed models. In a subsequent study, McPherson and McCormick (2006) found adequate fit between their data and a revised model, χ^2 (364, $N = 446$) = 1837.78, $p < .01$, adjusted goodness of fit index (AGFI) = .93, and root mean square error of approximation (RMSEA) = .08. In this model, self-efficacy was an endogenous, or mediating variable,

positioned between the variables cognitive strategy use, practice time, practice regulation, grade level, formal, and informal practice and the outcome variable performance achievement. To summarize, self-efficacy mediated all of the variables' relationships with performance achievement. The results of this study supported their previous findings of self-efficacy as a significant predictor of achievement in music performance. These studies by McPherson and McCormick examined the relationship between self-efficacy and music performance, but did not extend the investigation to include the sources of self-efficacy as did Lent and others (1994) and Fouad and Smith (1996). This omission created a need to extend this line of research to include the sources of self-efficacy.

Purpose

The primary purpose of this study was to develop a greater understanding of self-efficacy in music performance among secondary school music students. In order to accomplish that goal, however, I found it necessary to evaluate the effectiveness of the Music Performance Self-Efficacy Scale as a measure of self-efficacy. Through this evaluation, I was able to investigate the relationships between the sources of self-efficacy and the composite construct of self-efficacy. I was also able to compare scores among students in different music ensembles and grade levels to determine if secondary school music students attributed the same values to the sources of self-efficacy. In addition, I extended the investigation of self-efficacy to explore the relationship between self-efficacy in music performance and music aptitude. Both music aptitude and self-efficacy have been linked to music achievement in prior studies, but there was gap of knowledge regarding the relationship between the two constructs. Finally, results from this study

were used as evidence to determine whether the MPSES generated scores that were valid and reliable measures of self-efficacy in music performance.

Research Questions

1. Do the responses of secondary school music student to items on the Music Performance Self-Efficacy Scale adequately fit Bandura's theoretical model of self-efficacy?
2. How much influence do mastery experience, vicarious experience, verbal/social persuasion, and physiological state have on self-efficacy in music performance among secondary school music students?
3. Are the items on the Music Performance Self-Efficacy Scale invariant for middle and high school music students?
4. Are there differences in the value attributed to mastery experience, vicarious experience, verbal/social persuasion, and physiological state by band, chorus, and string orchestra students?
5. Are there differences in the value attributed to mastery experience, vicarious experience, verbal/social persuasion, and physiological state by middle and high school music students?
6. To what extent does music aptitude predict self-efficacy in music performance?
7. To what extent does evidence support the use of the Music Performance Self-Efficacy Scale as a measure of self-efficacy in music performance?

Significance of the Study

This study was significant because it contributed to the fundamental understandings of self-efficacy in music performance. Previously, self-efficacy had been found to be the best predictor of achievement in music performance (McCormick & McPherson, 2003; McPherson & McCormick, 2000, 2006) and the primary mediating variable between anxiety and performance outcomes (Craske & Craig, 1984). As an extension of this previous work, the current study investigated the sources of self-efficacy and their contribution to the development of self-efficacy beliefs. Scales to measure the sources of self-efficacy have been constructed in other domains (Lent, Lopez, & Bieschke, 1991; Matsui, Matsui, & Ohnishi, 1990; Usher & Pajares, 2006, 2009), but not in music performance. This study analyzed data produced by the Music Performance Self-Efficacy Scale and other measures to provide information about the sources of self-efficacy in music performance, as well as evidence to determine the validity and reliability of the scores generated by the MPSES.

Operational Definition of Terms

Independent variables.

Grade Level – This variable refers to the grouping of grade levels in secondary schools. There are two levels, middle school and high school. Middle school includes grades 6-8, and high school includes grades 9-12.

Ensemble – This variable refers to the ensemble in which the study participants perform music. There are three levels, band, chorus, and string orchestra. For the purpose of this study, participants were included in a level based on the class in which they participated in the study. Participants may have participated in multiple ensembles, but

they were only allowed to complete the questionnaires one time. Only students in these three types of ensembles were included. Students in ensembles such as guitar or steel drums were excluded.

Dependent variables.

Mastery Experience – This variable refers to the composite score of items from the mastery experience section of the Music Performance Self-Efficacy Scale (see Appendix B). There are eight items in this section (#1, 4, 6, 8, 10, 12, 14, 16). Mastery experiences are memories of an individual's past successes or failures in an activity.

Vicarious Experience – This variable refers to the composite score of items from the vicarious experience section of the Music Performance Self-Efficacy Scale. There are five items in this section (#2, 5, 11, 18, 20). Vicarious experiences are formed by observing others with similar characteristics to an individual engage in activities, and then predicting success or failure of the individual in the activity based upon those observations.

Verbal/Social Persuasion – This variable refers to the composite score of items from the verbal/social persuasion section of the Music Performance Self-Efficacy Scale. There are six items in this section (#3, 7, 9, 13, 21, 22). Verbal/social persuasions are based on the judgments and opinions of others and the influence of those judgments and opinions on an individual's self-beliefs.

Physiological State – This variable refers to the composite score of items from the physiological state section of the Music Performance Self-Efficacy Scale. There are five items in this section (#15, 17, 19, 23, 24). Physiological states such as anxiousness or nervousness may indicate feelings of vulnerability or ineptitude.

Music Performance Self-Efficacy – This variable refers to the composite score of all items from the Music Performance Self-Efficacy Scale. There are 24 items. Consistent with Bandura's theory, the composite score is an aggregate of scores from each section, or source of self-efficacy. Music performance self-efficacy refers to an individual's belief in his or her ability to successfully perform music.

Mathematics Self-Efficacy – This variable refers to the composite score of items from the Sources of Middle School Mathematics Self-Efficacy Scale (Usher & Pajares, 2009) (see Appendix D). There are 24 items. Mathematics self-efficacy refers to an individual's belief in his or her ability to successfully complete mathematics tasks.

Self-Esteem of Musical Ability – This variable refers to the composite score of items from the Self-Esteem of Music Ability (Schmitt, 1979) (see Appendix E). There are 43 items. Self-esteem of music ability refers to the personal estimation of an individual towards his or her capacity to perform music tasks and obtain musical competence.

Music aptitude – This variable refers to the composite score of items from the Advanced Measures of Music Audiation (Gordon, 1989) (see Appendix F). There are 30 items. Music aptitude refers to an individual's innate potential to learn music.

Teacher ratings – Teachers used their perceptions of student behavior in the classroom to estimate the student's self-efficacy for music performance. They rated their students on a 5-point Likert-type scale (1 = very low, 5 = very high) (see Appendix G).

Delimitations

The generalization of results from this study is limited to the study participants. This limitation was determined by the use of a convenience sample. Convenience sampling was necessary because the participants were nested within specific music

ensembles and it was not feasible to gather a random sample of secondary school music students. The diversity of the participants' geographic locations and schools, however, enhanced the external validity of the findings. In addition, there were threats to the ecological validity of the results. These threats included the influence of testing conditions and the Hawthorne effect. Since the questionnaires were administered online at various sites, testing conditions could not be controlled. These conditions included the speed of the testing computers, the quality of the audio playback, and the comfort of the testing environment. Also, participants were aware that they were participating in a research study and their self-reported responses may be influenced by their desire to please their music teacher or the researcher. The results must be interpreted with caution.

Limitations

Because of the descriptive nature and one-shot design of this study, the 12 extraneous variables affecting internal validity in experimental research (Gall, Gall, & Borg, 2007) did not apply: history, maturation, testing, instrumentation, statistical regression, differential selection, experimental mortality, selection-maturation interaction, experimental treatment diffusion, compensatory rivalry by the control group, compensatory equalization of treatments, and resentful demoralization of the control group. Nevertheless, experimenter bias was one extraneous variable that had the potential to influence the internal validity of the study. To control experimenter bias, the principal investigator had no direct contact with the participants and the participants were given no information regarding the hypotheses or anticipated outcomes of the study. Although this technique was used to control one extraneous variable's influence on internal validity, the

study's cross-sectional design (i.e., different age groups) limited the degree to which the influences of other extraneous variables could be determined and controlled.

Chapter 2: Review of the Literature

This chapter is a review of the literature related to the self-perception of self-efficacy and its role in music performance. For clarity, it consists of three sections: self-efficacy, self-efficacy scales, and self-efficacy in music education. In the self-efficacy section, I describe the psychological construct of self-efficacy from the perspective of Bandura's social cognitive theory. This theory provided the theoretical foundation for development of the Music Performance Self-Efficacy Scale and provided the basis for interpretation of scale results. In the self-efficacy scales section, I summarize research related to the development and validation of self-efficacy scales. I discuss findings in other studies and highlight important measurement and validation issues that may be problematic in this investigation. Finally, in the self-efficacy in music education section, I review the findings of other music education researchers regarding self-efficacy in music performance. This review includes a brief history of studies related to self-perceptions of ability in music performance and the scales developed to test them.

Self-Efficacy

Bandura (1977) introduced the term self-efficacy within the context of social learning. He later incorporated this psychological construct into his social cognitive theory. In his theory, Bandura (1986) proposed a model of learning in which cognition, behavior, and environment influenced one another. Individuals processed information from each of these elements to form thoughts and determine behavior. Self-efficacy functioned as self-referent thought within this model mediating the relationship between

cognition and behavior. Bandura stated, "Perceived self-efficacy is defined as people's judgments of their capabilities to organize and execute courses of actions required to attain designated types of performances" (1986, p. 391). From this perspective, individuals hold high levels of self-efficacy when they believe the level of their skills meets or exceeds the demands of the task. He indicated that self-efficacy was not a global trait, but tied to specific tasks and situations. In sports for example, an individual may have a strong sense of self-efficacy for playing golf and a weak sense of self-efficacy for playing baseball. He also proposed that self-efficacy beliefs could change over time. Berry and West (1993) summarized this characteristic of self-efficacy stating, "it is dynamic and malleable, subject to changes in task demands, situational determinants, social context, and individual development" (p. 353). Due to this flexible quality, self-efficacy has become a topic of interest for educational researchers (Pajares, 1996; Schunk 1983, 1990; Usher & Pajares, 2006, 2009).

As mentioned above, Bandura (1977, 1986) identified four sources of information that contributed to the development of self-efficacy beliefs. Those sources were: (a) enactive mastery experience, (b) vicarious experience, (c) verbal/social persuasion, and (d) physiological state. He believed mastery experience had the strongest influence on self-efficacy beliefs and was based on an individual's past successes or failures in an activity. Vicarious experiences had less influence than mastery experiences and were the result of observing others similar to an individual engage in activities and then predicting outcomes based upon those observations. Verbal/social persuasions were the judgments and opinions of others and their influence on an individual's decision making process. Physiological states referred to levels of physical arousal and emotional mood.

Uncomfortable states such as anxiousness or nervousness indicated underlying feelings of vulnerability and ineptitude that undermined the other sources. Negative feelings reduced an individual's belief in his or her ability to complete a task successfully. These four sources contributed to the development of a composite self-efficacy belief.

Initially, studies of self-efficacy were conducted in the field of psychology. Psychologists measured self-efficacy to evaluate the effectiveness of psychotherapy on patients' self-perceptions of their abilities to confront phobias (Bandura, 1982). These therapies were designed to improve patients' self-efficacy, and as self-efficacy increased, the effects of the phobia decreased.

Since then, studies of self-efficacy have been conducted in the field of education. Researchers have investigated the relationship between self-efficacy and achievement in several subject areas. Usher and Pajares (2006) identified a positive relationship between academic self-efficacy and reading grade ($r = .32$, $p < .001$). Other researchers investigated the relationship between self-efficacy and mathematics achievement (Lent, Lopez, & Bieschke, 1991; Lopez & Lent, 1992; Matsui, Matsui, & Ohnishi, 1990; Usher & Pajares, 2009). In one study, Lopez and Lent (1992) found self-efficacy scores in mathematics were positively correlated with mathematics grade ($r = .50$, $p < .01$). In science, Joo, Bong, and Choi (2000) identified a significant relationship between academic self-efficacy and scores on a science achievement test ($r = .46$, $p < .001$). The results of these studies confirmed Bandura's premise that self-efficacy was related to achievement. The relationship between self-efficacy and achievement, however, may not be the same for all students. Pajares and Valiante (1999) found a stronger relationship between writing self-efficacy and grade point average among middle school girls ($r = .38$,

$p < .001$) than boys ($r = .21$, $p < .001$). Differences in gender and other characteristics may determine differences in an individual's self-efficacy beliefs.

Education researchers have also examined the relationships between the sources of self-efficacy and other variables. Usher and Pajares (2006) adapted Lent, Lopez, and Bieschke's (1991) Sources of Math Efficacy Scale to measure the sources of academic self-efficacy among middle school students and correlated those scores with results from the academic self-efficacy section of Bandura's Children's Multidimensional Self-Efficacy Scale (CMSES). Within a convenience sample of 263 6th grade participants, they found each source of self-efficacy to be significantly related ($p < .001$) to academic self-efficacy (mastery experience $r = .57$, vicarious experience $r = .39$, verbal/social persuasion $r = .45$, and physiological state $r = -.39$). In the same study, Usher and Pajares correlated the sources of self-efficacy scores with results from the self-regulation section of Bandura's CMSES. Once again, they found each source of self-efficacy to be significantly related ($p < .001$) to self-efficacy for self-regulation (mastery experience $r = .63$, vicarious experience $r = .44$, verbal/social persuasion $r = .51$, and physiological state $r = -.44$). It is important to note the similarities of the relationships in the two sets of results. In both cases, mastery experience exhibited the strongest relationship with composite construct. These results verified Bandura's proposal that mastery experiences had the most influence among the four sources of self-efficacy. It is important to note that the negative correlations between physiological state and academic self-efficacy and between physiological state and self-regulation did not contradict the findings of the other self-efficacy sources. Rather, the negative correlations reflected the reverse coding, or wording, of the items. Items such as "School work makes me nervous and

uncomfortable" (Usher & Pajares, 2006, p. 131), elicited responses that were opposite to the responses for items reflecting the other sources of self-efficacy.

The order in which information from the four sources is received has been another area of interest. Wise and Trunnell (2001) examined the influence of three sources of self-efficacy on participants' bench-press efficacy. Consistent with Bandura's theory, they found mastery experience to be a better predictor of bench-press efficacy than vicarious experience ($t = 1.82$, $p = .04$). They also found that vicarious experience was a better predictor than verbal/social persuasion ($t = 1.81$, $p = .04$). In a unique approach, Wise and Trunnell examined the influence of the sources by the order in which they were presented to participants. They found that verbal/social persuasion was more effective when following a mastery experience than when following a vicarious experience ($t = -3.76$, $p = .003$). Participants placed more value on verbal/social persuasion after they had accomplished a task, as opposed to watching someone else perform a task. This finding may have implications for music teachers and the delivery of positive verbal reinforcement during rehearsals. One criticism of this study, however, is the author's failure to control Type I error. The authors should have adjusted the significance level to avoid increasing probability of Type I error since multiple *t*-tests were conducted using the same participants.

To summarize, I used Bandura's social cognitive theory as the theoretical framework for this study. In his theory, Bandura identified self-efficacy as the primary determinant of human thought and action. He proposed that four sources of information (i.e., mastery experience, vicarious experience, verbal/social persuasion, and physiological state) contributed to the development of self-efficacy beliefs. Education

researchers have confirmed relationships between self-efficacy and various forms of achievement such as reading grade (Usher & Pajares, 2006), mathematics grade (Lopez & Lent, 1992), and science test scores (Joo et al., 2000). Researchers have also found the magnitude of the correlations between the sources of self-efficacy and academic self-efficacy and between the sources of self-efficacy and self-regulation (Usher & Pajares, 2006) to be consistent with Bandura's proposed model. Mastery experience had the strongest correlation with the other variables, followed by verbal/social persuasion, physiological state, and vicarious experience. These findings provided a foundation of supporting evidence on which to build this investigation of self-efficacy in music performance.

Self-Efficacy Scales

The following section provides information about the construction of self-efficacy scales, an overview of previously constructed scales, and an introduction to the issues related to interpreting results from this type of scale. The section is divided into eight subsections: (a) guidelines for constructing self-efficacy scales, (b) general scales of self-efficacy, (c) self-efficacy scales in education, (d) sources of self-efficacy scales, (e) validity, (f) reliability, (g) response formats, and (h) age and grade differences. I used the information presented in these subsections to determine direction and establish a context for the current study. Results from other scales were used as a resource with which to compare findings from the current study. These comparisons offered insights into the functioning of the items on the Music Performance Self-Efficacy Scale and examination of self-efficacy in music performance.

Guidelines for constructing self-efficacy scales.

Bandura (2006) provided guidelines for constructing self-efficacy scales. In these guidelines, self-efficacy was not characterized as a global trait, but rather a "set of self-beliefs linked to distinct realms of functioning" (p. 307). He believed self-efficacy beliefs were specific to knowledge domains and physical activities. He also suggested that scales of self-efficacy should accurately reflect the construct of self-efficacy and differentiate it from similar constructs such as self-esteem, self-confidence, and outcome expectancies. Bandura recommended that scale items cover various levels of task demands by including differences in generality, strength, and level. Generality referred to variations in the types of activities and contexts, strength addressed the intensity of self-efficacy in relation to the task, and level reflected the inherent difficulty of the task. Once the items were constructed, he advised scale developers to analyze the items for their ability to differentiate between participants and be aware that items reflecting the same domain of efficacy should correlate with each other and with the total score. He also recommended the use of factor analysis as a technique for evaluating the homogeneity of the items.

General scales of self-efficacy.

Researchers have developed scales of self-efficacy based on Bandura's theoretical concepts. Sherer and others (1982) developed one of the first self-efficacy scales. During the scale development process, participants responded on a 14-point Likert-type scale (1-strongly disagree, 14-strongly agree) to 36 initial items. Using exploratory factor analysis, only 23 of those items loaded on two factors with loadings greater than .40. Consequently, 13 items were discarded and the authors labeled the two factors general self-efficacy and social self-efficacy. The 17 items that loaded on the first factor

constituted the general self-efficacy section of the scale and accounted for 26.5% of the total variance. The remaining six items loaded on the second factor and comprised the social self-efficacy section of the scale and accounted for 8.5% of the total variance. The internal consistency for each section was $\alpha = .86$ and $\alpha = .71$ respectively. The scale was developed for use in the field of psychology to provide therapists with a measure of therapeutic progress.

Bandura developed two general self-efficacy scales, the Children's Perceived Self-Efficacy Scales (1990, as cited in Pastorelli, Caprara, Barbaranelli, Rola, Rozsa, & Bandura, 2001) and the Multidimensional Scales of Perceived Self-Efficacy (1990, as cited in Choi, Fuqua, & Griffin, 2001). The Children's Perceived Self-Efficacy Scale (CPSES) measured self-efficacy in seven domains: (a) academic achievement, (b) self-regulated learning, (c) leisure and extracurricular activities, (d) self-regulatory efficacy, (e) maintaining social relationships, (f) self-assertive efficacy, and (g) meeting others' expectations. Given the subject area of the current study, it should be noted that there was one item in the leisure and extracurricular activities that asked participants to rate their ability to learn music skills.

In an international study with 10-15 year old participants from Italy, Hungary, and Poland, Pastorelli and others (2001) used exploratory factor analysis to identify the factor structure among participants from each country on the CPSES. Scores from each country loaded on the same three factors which the researchers labeled as academic self-efficacy, social self-efficacy, and self-regulatory learning. There were significant differences, however, among the independent variables of gender and country. Compared with boys, girls reported higher levels of academic self-efficacy, $F(1, 1179) = 30.165$, $p < .001$, and

higher levels of self-regulatory learning, $F(1, 1179) = 20.688$, $p < .001$. In a comparison by country, Italian children demonstrated higher academic self-efficacy and social self-efficacy than Hungarian children ($p < .05$). These results supported beliefs that self-efficacy varied by gender and also national origin. Pastorelli and others suggested these disparities were evidence of social and cultural differences among the participants.

Bandura's Multidimensional Scales of Perceived Self-Efficacy (1990, as cited in Choi et al., 2001) was tested with undergraduate students. The Multidimensional Scales of Perceived Self-Efficacy was similar to the Children's Perceived Self-Efficacy Scale except that it measured nine domains instead of seven. The two additional domains were enlisting social resources and enlisting parent and community support. The internal consistency of the scores for each of the nine domains ranged from .63 for enlisting social resources to .86 for self-regulated learning. The researchers, however, were unable to load the items on the same three factors of academic self-efficacy, social self-efficacy, and self-regulatory learning as in Pastorelli and others (1990) analysis of the CPSES. Further study into the differences between these scales may be warranted.

Self-efficacy scales in education.

Researchers have created scales to measure self-efficacy in a variety of educational domains. Pintrich, Smith, Garcia, and McKeachie (1991) created the Motivated Strategies for Learning Questionnaire (MSLQ) and included a section that measured academic self-efficacy. The section consisted of eight items that addressed self-efficacy in learning. The items were general in nature and could be adapted for use in a variety of contexts and subject domains. The internal consistency among the items was high ($\alpha = .93$) and the authors recommended using this section as a separate measure of

self-efficacy. Following their advice, the section was adapted to measure music self-efficacy among middle school students in traditional and technology-enhanced music learning environments (Zelenak, 2009). Although no difference was found between the two groups, the section functioned well as a measurement instrument. In addition to the self-efficacy section of the MSLQ, Usher and Pajares (2006) created their own scale of academic self-efficacy. Their scale consisted of 24 items and was an adaptation of an earlier work by Lent, Lopez, and Bieschke (1991). The primary difference between Usher and Pajares's scale and the self-efficacy section of Pintrich and others' MSLQ was that Usher and Pajares's scale evaluated Bandura's four sources of academic self-efficacy, while the MSLQ section measured the composite construct of academic self-efficacy.

Researchers have developed scales of self-efficacy for specific subjects. Aydin and Uzuntiryaki (2009) created the High School Chemistry Self-Efficacy Scale. Their 16-item test loaded successfully on two latent variables, chemistry self-efficacy for cognitive skills and chemistry laboratory knowledge. Scores generated by each section were found to be internally consistent (α = .90 and α = .92) and the items were unbiased for public and private schools. Also in the subject of science, Dalgety, Coll, and Jones (2003) assembled the Chemistry Attitudes and Experiences Questionnaire which included a section dedicated to self-efficacy in chemistry. In other subjects, Betz and Hackett (1983) developed a mathematics self-efficacy scale for undergraduate students, and Pajares (2007) designed a scale to assess the writing of students in grades 4-11. These scales have been psychometrically tested and were based on Bandura's theoretical framework.

Sources of self-efficacy scales.

The relationships between the sources of self-efficacy and the composite construct of self-efficacy have been examined in a variety of subject areas. Although most of the work has been done in the field of mathematics (Lent, Lopez, & Bieschke, 1991; Lopez & Lent, 1992; Matsui, Matsui, & Ohnishi, 1990; Usher & Pajares, 2009), scales have also been developed to investigate the relationships in academic self-efficacy (Usher & Pajares, 2006), jazz experience (Wehr-Flowers, 2007), and orchestra performance experience (Hendricks, 2009). Participants in these investigations have included middle school students, high school students, and college undergraduates. The number of items used in the scales has ranged from 14 to 87. The response formats for most scales were based on Likert-type responses (Table 1).

Table 1
General Descriptions of Sources of Self-Efficacy Scales

Scale/Subject area	Participants	Item Information	Response Format
"no title" / mathematics Matsui, Matsui, & Ohnishi (1990)	undergraduates (*N* = 163)	16 items mastery experience: high school math grade vicarious experience, verbal/social persuasion, emotional arousal: 5 researcher developed items each.	5-point Likert-type scale (1-lowest, 5-highest)
"no title" / mathematics Lent, Lopez, & Bieschke (1991)	undergraduates (*N* = 138)	40 items mastery experience, vicarious experience, verbal/social persuasion: 10 researcher developed items each emotional arousal: 10 items from Fennema-Sherman Math Anxiety Scale	5-point Likert-type scale (1-lowest, 5-highest)
Sources of Math Efficacy Scale / mathematics Lopez & Lent (1992) adapted from Lent et al. (1991)	high school students (*N* = 50)	40 items (see Lent et al., 1991)	5-point Likert-type scale (1-lowest, 5-highest)
Sources of Self-Efficacy Scale / academic Usher & Pajares (2006) adapted from Lent et al. (1991)	6th grade students (*N* = 263)	24 items mastery experience – 6 items vicarious experience – 6 items (3 peer, 3 adult) social persuasion – 5 items	6-point Likert-type scale (1-lowest, 6-highest)

Table 1 (continued)
General Description of Sources of Self-Efficacy Scales

Scale	Participants	Item Information	Response Format
Sources of Middle School Mathematics Self-Efficacy Scale/ mathematics Usher & Pajares (2009)	students in grades 6, 7, and 8 (*N* = 803)	24 items mastery experience vicarious experience verbal/social persuasion physiological state – 6 researcher developed items each	6-point Likert-type scale (1-lowest, 6-highest)
Jazz Experience Survey / music Wehr-Flowers (2007)	undergraduates (*N* = 281)	sources of self-efficacy: 87 of 161 items mastery experience: 28 items based on Midgley et al. (2000), vicarious experience: 17 items, 11 items researcher designed, 6 items from Sawyer & Hollis (2005) verbal/social persuasion: 19 items, 5 items researcher designed, 6 items from Marx & Stapel (2004), 8 items from Pascarella (1996), physiological state: 23 items from Friedman & Brendas-Jacob (1997)	10 unit intervals on 0-100 scale
Sources of Self-Efficacy: Post Rehearsal Questionnaire / music Hendricks (2009)	high school students (*N* = 157)	14 items mastery experiences – 4 items vicarious experience – 3 items verbal persuasion – 3 items physiological state – 4 items	11-point scale (-5 to +5).

Researchers have been able to develop items that consistently measure three of the four self-efficacy sources: mastery experience, verbal/social persuasion, and physiological state. Weaknesses have appeared, however, in constructing items to measure vicarious experience. Some researchers (Lent, Lopez, Brown, & Gore, 1996; Usher & Pajares 2009) have addressed this problem by differentiating vicarious experiences into "peer" and "adult" categories. Although the internal consistency values improved, Lent and others (1996) found no significant difference in the chi-square value, or degree of fit, between the four and five factor models in a study with high school and college students. This result implied that there may be little advantage in separating vicarious experience into two categories. In a comparison of the internal consistencies, vicarious experience demonstrated the lowest level among all reviewed research studies (Table 2).

Table 2
Internal Consistencies (Cronbach's α) of Sources of Self-Efficacy Scales

Source	Authors, (Year), and Subject				
	Matsui, Matui, & Ohnishi (1990) Math (N =163)	Lent, Lopez, & Ohnishi (1991) Math (N = 138)	Lopez & Lent (1992) Math (N = 50)	Usher & Pajares (2006) Academics (N = 263)	Usher & Pajares (2009) Math (N = 803)
Mastery Experience	--[a]	.86	.82	.86	.88
Vicarious Experience	.69	.56	.59	.68	.84
Verbal Persuasion	.91	.74	.74	.82	.88
Physiolog. State	.84	.90	.90	.84	.87

[a.] Used high school math grade.

In an attempt to investigate vicarious experience further, Usher and Pajares (2006) conducted an exploratory factor analysis of data generated by 6th grade students on a 24-item sources of self-efficacy scale. The analysis identified five factors. Factor 1 consisted of six items reflecting mastery experiences with loadings ranging from .54 to .83, factor 2 consisted of seven items reflecting physiological state with loadings ranging from .45 to .82, factor 3 consisted of five items reflecting social persuasion with loadings ranging from .41 to .84, and factors 4 and 5 consisted of vicarious experiences that were differentiated between peer (three items) or adult (three items) models. Loadings for the peer items ranged from .41 to .62, and loadings for the adult items ranged from .56 to .62. Usher and Pajares, however, found the internal consistency of the peer items unsatisfactory (α = .59) and decided to use only the items associated with adults (r = .72). Other studies have reported similar low levels of internal consistency for vicarious experiences and most continued to use the four factor model (Lent, Lopez, & Bieschke, 1991; Lopez & Lent, 1992; Matsui, Matsui, & Ohnishi, 1990). Developing items that produce internally consistent data for vicarious experience continues to be a challenge for researchers. To promote parsimony, the current study used four sources of self-efficacy and did not split vicarious experience into peer and adult categories. The issues surrounding vicarious experiences are reminders that self-efficacy is a reflection of personal interpretations of events rather than factual descriptions. Students interpret actions by peers and adults based on their social expectations and cultural norms.

Correlations between the sources of self-efficacy and general measures of self-efficacy have indicated a consistent rank order in the relationships between the sources and the general self-efficacy measure (Table 3). Consistent with Bandura's theory,

mastery experience demonstrated the strongest relationship with self-efficacy in all studies. Verbal/social persuasion exhibited the next strongest relationship followed by physiological state and vicarious experience. Lopez and Lent (1992) provided the only exception with math self-efficacy showing a slightly stronger correlation with vicarious experience (r = .26) than with physiological state (r = -.23). Correlations for physiological state are easily misinterpreted because the positive or negative wording of the items influences the positive or negative direction of the correlation.

Table 3
Correlations of Sources of Self-Efficacy Scores with Subject Self-Efficacy Scores

Source	Authors, (Year), and Subject				
	Matsui, Matsui, & Ohnishi (1990)[a] Math (N=163)	Lent, Lopez, & Biescke (1991)[a] Math (N = 138)	Lopez & Lent (1992)[a] Math (N = 50)	Usher & Pajares (2006)[b] Academic (N = 263)	Usher & Pajares (2009)[c] Math (N = 803)
Mastery Experience	.42	.63	.57	.57	.77
Vicarious Experience	.19	.15	.26	.39	.44
Verbal/Social Persuasion	.37	.54	.34	.45	.61
Physiological State	-.27	.49	-.23	-.39	-.55

[a] Correlated with math self-efficacy scale
[b] Correlated with academic self-efficacy scale
[c] Correlated with math grade self-efficacy

It is also important to note differences in the magnitude of the relationships for middle school, high school, and collegiate populations. Although the rank order of sources was consistent, the magnitude of the relationships differed by grade level. For

example, the relationship between mastery experience and general math self-efficacy was $r = .63$ for college undergraduates (Lent, Lopez, & Biescke, 1991) and the relationship was $r = .57$ for high school students (Lent & Lopez, 1992). This difference suggested that there were differences in the strength of the relationships based on age. These findings provided an rationale for investigating age differences in the current study.

Although correlations provided some information about the relationships between self-efficacy and other variables, regression analysis enabled further insight into the contributions of the sources (Table 4). A comparison of standardized regression coefficients across studies confirmed that mastery experience contributes a larger proportion of the increase in self-efficacy, or criterion variable, than the other sources. In turn, mastery experience also accounted for a larger percentage of the variance in scores. The other sources differed across studies in the relative strengths of their regression coefficients indicating differences in their abilities to predict outcomes and account for variance. Standardized regression coefficients were particularly useful in this context because they could be compared across studies.

Table 4
Comparison of Standardized Regression Coefficients among Self-Efficacy Scales

Source	Authors, (Year), and Subject			
	Lent, Lopez & Biescke (1991)[a] Math (N = 138)	Lent & Lopez (1992)[a] Math (N = 50)	Usher & Pajares (2006)[b] Academic (N = 263)	Usher & Pajares (2009)[c] Math (N = 803)
Mastery Experience	.36	.57	.34	.64
Vicarious Experience	.01	.09	.00	.03
Verbal Persuasion	.16	.00	.16	.08
Physiological State	.02	.32	-.62	-.09

[a] Regressed on mathematics self-efficacy scale
[b] Regressed on academic self-efficacy scale
[c] Regressed on math grade self-efficacy

Confirmatory factor analysis (CFA) was used in several studies to determine whether or not the data in each study fit Bandura's proposed model. Put simply, the CFA compares the covariance matrix of the collected data with the covariance matrix of the proposed model. Initially, the χ^2 statistic was used to test and interpret the comparison, but deemed unreliable due to its inability to adjust for larger sample sizes. Consequently, researchers have developed other indices that are more reliable with large sample sizes. Commonly used indices include the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). Under most conditions, researchers have found that the CFI values $\geq$ .95, RMSEA values $\leq$ .06, and SRMR values $\leq$ .08 indicate an acceptable likelihood that the matrices are similar (Hu & Bentler, 1998).

There has been a variety of findings associated with the use of CFA in the evaluation of sources of self-efficacy scales. Usher and Pajares (2009) were able to find an acceptable level of fit between data produced by their Sources of Middle School Mathematics Self-Efficacy Scale and Bandura's proposed sources of self-efficacy model, CFI = .96, RMSEA = .04, and SRMR = .04. These index values satisfied Hu and Bentler's (1998) benchmarks for adequate fit. Some studies in music education, however, have not been able to attain adequate fit. Wehr-Flowers (2007) was unable to make the data in her study converge with Bandura's proposed model. This lack of convergence may have been a symptom of problems with the items, the model design, or numerous other issues. As an alternative to treating the sources as 1st order factors and self-efficacy as a 2nd order factor, Wehr-Flowers chose to treat them as independent variables in her analysis. The problem experienced by Wehr-Flowers exemplified the need for a valid and reliable self-efficacy scale in music performance. In another music study, Hendricks (2009) used exploratory factor analysis to examine the 14 items in her proposed sources of self-efficacy scale. The 14 items loaded on only two factors, rather than the four recommended by Bandura. While Usher and Pajares's (2009) study indicated the potential for creating a scale fitting Bandura's framework, Wehr-Flowers (2007) and Hendricks (2009) brought to light problems music researchers have encountered in constructing self-efficacy scales for music performance. Music performance is a unique activity requiring the simultaneous coordination of several types of skills. A scale that measures self-efficacy in music performance must reflect the specialized demands of the activity.

Validity.

Determining a scale's ability to generate scores that are valid is an essential part of the scale construction process. Messick (1989) suggested that validity was not a characteristic of the scores themselves, but an evaluation of the inferences and actions that result from those scores based on various sources of empirical and theoretical evidence. His perspective is consistent with the standards for validity recommended by the American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education (1999). In an exemplary study, Usher and Pajares (2009) used multiple methods to establish evidence with which to assess the validity of results from their Sources of Middle School Mathematics Self-Efficacy Scale. To establish evidence of test content in phase 1 of the study, they formed a focus group of 23 6th grade students, two math teachers, a parent of middle school students, and a school principal to evaluate the scale items. After preliminary testing and revisions, the items were then sent to three experts of social cognitive theory (A. Bandura, B. J. Zimmerman, and D. H. Schunk) for further content evaluation. To establish evidence of response processes, participants responded to items written as first-person statements and rated their self-perceptions on a 6-point Likert-type scale. This response style was consistent with Bandura's belief that self-efficacy was based on self-perception and interpretation of experiences rather than achievements and measurable outcomes.

In phase 2 of the study, Usher and Pajares conducted an exploratory factor analysis to verify evidence of internal structure. They used EFA to identify underlying factors and evaluate the quality of items based upon the strength of their loadings. In

phase 3, or the final phase of the study, they used confirmatory factor analysis to verify the fit of the data to the proposed model. They found that the data adequately fit the proposed model, χ^2 (246, N = 803) = 601.21, p < .001, CFI = .96, RMSEA = .04, SRMR = .04. They also constructed four multiple regression equations using four measures of self-efficacy as criterion variables, one in each equation. The criterion variables were grade self-efficacy, mathematics skills self-efficacy, courses self-efficacy, and self-efficacy for self-regulated learning. The four sources of self-efficacy were used as predictor variables in those equations. The results verified Bandura's assertion that mastery experience was the strongest and most consistent predictor of self-efficacy. This finding was then used as evidence of criterion validity.

To indicate evidence of relationships with other variables, Usher and Pajares compared several variables to establish convergent and discriminant validity. They stated, "Convergent validity was supported by the strong association between the sources, self-efficacy, related motivation constructs, and achievement" (p. 97). They were also able to demonstrate the discriminate capability of their scale by finding only modest correlations between the sources of mathematics self-efficacy and the unrelated construct of performance goal orientation (r = .02 to r = .34). Usher and Pajares did not provide evidence based on the consequences of testing, but did provide the following recommendations for future use of their scale: (a) the relationships between the sources and self-efficacy would differ depending upon the measures used, (b) the relationships should not be generalized to other settings and contexts, and (c) reporting of the relationships should include descriptions of the participants. Usher and Pajares's study provided a model for establishing evidence of validity in the current study.

Reliability.

Along with validity, measurement integrity depends upon the reliability of the scores. Traditionally, reliability has been established in three ways: (a) comparison of parallel test forms, (b) test-retest, and (c) internal consistency. In their study, Usher and Pajares (2009) calculated only the internal consistency of their scores for each source of self-efficacy. The internal consistencies of the scores in each section were: mastery experience α = .88, vicarious experience α = .84, verbal/social persuasion α = .88, and physiological state α = .87. All of these values were considered acceptable in general studies for being above the .80 threshold recommended by Nunnally (1967, as cited in Henson, 2001).

Response formats.

Studies of self-efficacy have used a variety of response formats. Some studies used a 5-point Likert-type response format (Lent et al., 1991; Lopez & Lent, 1992; Matsui et al., 1990). Other studies used a 6-point Likert-type format (Usher & Pajares, 2006; Usher & Pajares, 2009). Bandura (2006) recommended the use of a 0-100 range with participants rating themselves in 10 unit intervals resulting in an 10-point Likert-type scale. Usher and Pajares (2009) allowed participants to select any number between 0 and 100 in responding to items on the Mathematics Skills Self-Efficacy test which they used to validate their Sources of Middle School Mathematics Self-Efficacy Scale. As a result, they found the internal consistency of the scores on the skills test to be higher (α = .95) than the internal consistency of the other measures in the same study that used a 6-point Likert-type scale.

In a prior study, Pajares, Hartley, and Valiante (2001) compared scores from tests using a 6-point Likert-type format and a 0-100 incremental format and found that the 0-100 format was psychometrically stronger than the Likert-type format. They stated, “The fine-grained discrimination of the 0-100 scale provided an assessment of self-efficacy that was not only more strongly related to the achievement indexes with which it was compared but also predictive of achievement in a regression model” (p. 220). They also suggested that the 0-100 scale was not difficult for children to conceptualize since it follows the same scoring format used to formulate their grades and score their assignments. From a psychometric perspective, Byrne (2005) advised against using ordinal level data for confirmatory factor analysis and advocated for the use of interval level data. There are differing opinions, however, on whether 5- and 6-point Likert scales produce ordinal or interval level data. A discussion of this topic is beyond the scope of this paper. The 1-100 response format seemed to be the better choice for the current study because it utilized interval level data and was capable of achieving greater internal consistency.

Music researchers have used a variety of methods to measure music self-efficacy. One approach has been to simply pose questions to the participants. McPherson and McCormick (2000) asked the question "How good a musician do you think you are in comparison with other students of your own age?" (p. 33) to measure of general self-efficacy, and the question, "What result do you think you will get for your exam today?" (p. 34) to measure performance self-efficacy. Participants responded to the first question by choosing Poor, Below Average, Average, Above Average, or Outstanding, and responded to the second question by choosing Unsatisfactory below 65, Low pass 65-69,

High pass 70-74, Low merit 75-79, High merit 80-84, Low distinction 86-90, and High distinction 91-100. In a following study, McCormick and McPherson (2003) added an additional item along with the previous two, "I have fully mastered the requirements for today's examination" (p. 42). Participants were asked to rate themselves on a 1 ("not very true of me") to 7 ("very true of me") scale.

In a subsequent study, McPherson and McCormick (2006) moved away from the use of a few questions to evaluate self-efficacy and toward the development of a scale to measure self-efficacy. They extended their initial questions and used an 11-point Likert-type response format. Participants responded to the question stem of "How confident are you of your AMEB exam that you can perform correctly your:" which was then followed by (a) "technical work", (b) "pieces", and (c) "sight-reading" (p. 327) as separate responses. The stem of the second question asked "How confident are you for your AMEB exam that you can answer correctly each question in the:" which was followed by (a) "aural section" and (b) "general knowledge section." Participants responded to these five questions by selecting 0% to 100% in 10 unit intervals. This scale incorporated Bandura's (2006) recommendations to use multiple questions that reflect generality for the subject content by separating responses for the aural and general knowledge sections. It also allowed participants to use a response format (i.e., 0% to 100% in 10 unit intervals) that allowed for varying degrees of responses.

Another approach used by researchers has been to adapt an existing self-efficacy scale to the context of music education by incorporating appropriate music vocabulary. Nielsen (2004) compared self-efficacy beliefs among college music majors using an adaptation of the academic self-efficacy section from Pintrich and others' (1991)

Motivated Strategies for Learning Questionnaire. For example, he modified the item "I'm confident I can understand the most complex material presented by the instructor in this course" (Pintrich et al., 1991, p. 13) to "I am confident I can perform the most complex music I practice" (Nielsen, 2004, p. 422). If handled carefully, adaptation of a preexisting scale can be an easy and effective method for constructing a scale. Problems may lie, however, in whether the scale has the inherent flexibility to address the subtleties of different content domains.

In another study, Ritchie and Williamon (2007) adapted items from Sherer and others' (1982) General Self-Efficacy subscale to create the General Music Self-Efficacy Scale. From that scale, Ritchie and Williamon divided the 17 items into the two categories of learning and performing to create two additional scales. These scales were self-efficacy of music learning and self-efficacy of music performing. The internal consistency of responses from the general music self-efficacy scale was acceptable (α = .83) while the consistency of items on the learning and performing scales were questionable (learning scale α = .78, performing scale α = .68) according to Nunnally (1967, as cited in Henson 2001). Correlations between scores from the three scales of self-efficacy were moderate, ranging from r = .59 to r = .64, indicating the scales were measuring related, yet different constructs. Participants also compared their own abilities in relation to their peers on 22 music skills and attributes, and completed a questionnaire on self-regulated learning. Correlations between the participant's self-ratings and scores on the three measures were more diverse ranging from r = .30 to r = .63 and all were significant at the $p < .05$ level. Significant correlations were found between the self-

regulated learning scores and related items on the music learning and music performing scales.

Wehr-Flowers (2007) took an alternative approach to measuring musical self-efficacy. She constructed a detailed and lengthy scale to measure self-efficacy in jazz performance by combining items from several established scales with her own items. The Jazz Experience Survey consisted of 161 items. Eighty-seven of the items measured various dimensions of self-efficacy and the remaining items measured perceptions of jazz performance and gender bias. After several modifications, Wehr-Flowers was unable to get her data to converge with the proposed model using confirmatory factor analysis. Combining items from different sources and addressing too many factors at the same time may have been part of the convergence problem. Choi and others (2001) also reported convergence problems with Bandura's Multidimensional Scales of Perceived Self-Efficacy. That scale had 57 items and 9 factors.

Age and grade differences.

Researchers have found differences in self-efficacy related to age and grade level. Choi and others (2001) conducted an exploratory factor analysis of Bandura's Multidimensional Scales of Perceived Self-Efficacy. They anticipated results to be similar to Pastorelli and others (2001) earlier analysis of Bandura's Children's Perceived Self-Efficacy Scale in which the items loaded on three factors of academic self-efficacy, social self-efficacy, and self-regulatory learning since the scales were similar in content. They found, however, that the scores from the multidimensional scale would not load on the same three factors. This problem with the factor loadings may not reflect a difference in the scales, but rather the difference in the participants' ages. Pastorelli and others

studied participants in the 10-15 year age group, while Choi and others studied undergraduates with a mean age of 20.6 (SD = 3.60).

In another study, Zimmerman and Martinez-Pons (1990) examined the differences in verbal and mathematics self-efficacy among students in the 5th grade (10-11 years of age), 8th grade (13-14 years of age), and 11th grade (16-17 years of age). Using multivariate analysis of variance (MANOVA), they found a significant difference among scores on the dependent variables verbal self-efficacy and mathematics self-efficacy by grade level, $F_{mult}(4, 334) = 20.75$, $p < .01$, $R = .59$. Univariate follow-up procedures using analysis of variance (ANOVA) indicated significant differences among grade levels from the verbal and mathematics self-efficacy scores, $F(2, 168) = 29.82$, $p < .01$, $MS_e = 0.11$. *Post hoc* comparisons using the Newman-Keuls procedure found significantly higher levels of verbal self-efficacy among 11th graders than among 8th graders, and also among 8th graders when compared to 5th graders. In mathematics self-efficacy, 11th and 8th graders' scores were significantly higher than 5th graders, but no difference was found between 11th and 8th graders. These findings suggested that the degree of self-efficacy for verbal and mathematical tasks may be sensitive to age and grade level.

There has been great interest among researchers in measuring self-efficacy. Bandura provided several guidelines for constructing self-efficacy scales. These guidelines offered suggestions for constructing scale items, establishing a response format, and analyzing results. General self-efficacy scales from the field of psychology have been used as models to construct subject specific scales in education. Scales have been developed to measure the composite construct of self-efficacy and also the four sources. Patterns have emerged from the results of these scales supporting Bandura's

social cognitive theory. Mastery experience has been established as having the strongest influence on self-efficacy. Researchers have also identified problems measuring vicarious experience. Some have experimented with splitting the source into experiences watching peers and experience watching adults but have achieved limited success. Additional problems have been found in attempts to measure self-efficacy in music performance. Lack of convergence during analysis (Wehr-Flowers, 2007) and inability to load results onto Bandura's four sources (Hendricks, 2009) were two examples that may have been related to flaws in design. These examples made this researcher wary of the potential risks in developing a self-efficacy scale for music performance.

Self-Efficacy in Music Education

Self-perceptions of music ability have long been recognized as influencing music achievement. One type of self-perception is self-concept. Schmitt (1979) described self-concept as a set of organized and consistent attitudes about one's characteristics and abilities. Over 40 years ago, Greenberg (1970) attributed the lack of success among young male chorus singers to some type of emotional or psychological block. Greenberg stated, "continued out-of-tune singing and underachievement in music are primarily results of a low self-concept of one's ability to succeed in musical endeavors" (p. 57). In another study, Svengalis (1978) found self-concept of music among boys was positively correlated with their attitudes about music. The correlations ranged from $r = .39$ for boys with a low music self-concepts to $r = .74$ for boys with a high music self-concept. Austin and Vispoel (1998) used attribution theory to investigate music self-concept and music achievement. They found a moderate, yet significant, correlation between music self-concept and music achievement ($r = .50$, $p < .01$). More recently, Randles (2011)

discovered a negative relationship between grade level and music self-concept. As students' grade level increased, their self-perception of being a good musician decreased.

Another type of self-perception that has captured researchers' interest has been self-esteem. Hendricks (2009) defined self-esteem as an emotionally dependent evaluation of one's musical self-worth. Schmitt (1979) developed the Self-Esteem of Musical Ability (SEMA) to measure self-esteem among school aged children. She believed that children's self-esteem of music ability was the variable that accounted for discrepancies between their music aptitude and music achievement. In her study, she found that music self-esteem varied with the student's degree of success and involvement in music education programs. The more success students had in more music programs, the higher their levels of self-esteem. Randles (2010) identified the relatively stable nature of self-esteem as measured in a pre- and post-test assessment using Schmitt's SEMA. He learned, however, that students who engaged in music composition activities for 12 weeks demonstrated a significant increase in their self-perception of music ability ($r = .35$, $p < .01$). Engagement with music compositional activities was a stronger predictor of music self-esteem than involvement in music activities outside of school, performing in school ensembles, parental musical involvement, satisfaction level, and listening experiences.

As mentioned above, the self-perception of self-efficacy has also been found to influence music performance achievement. McPherson and McCormick (2000) found performance self-efficacy to be the best predictor of performance results among students aged 9-18 years. They stated, "students who display high self-efficacy expectations will be more likely to achieve in a difficult performance area, such as a formal music

examination, than their peers who display the same level of skill, but lower personal expectations" (p. 37). They also suggested that the ways in which students think about themselves, the task, and their performance were just as important as the time the students devoted to practicing their instrument. McCormick and McPherson (2003) verified the previous finding among a similar group of instrumentalists also aged 9-18 years. In this study, they extended their description of self-efficacy in music performance beyond performing on an instrument. They proposed that including judgments about the skill to perform in front of others was also an important part of music performance self-efficacy.

McCormick and McPherson (2003) investigated self-efficacy not just as a construct that influenced performance achievement, but also as a mediating variable between other constructs and performance achievement. Other constructs included cognitive strategy use, practice time, practice regulation, grade level, formal, and informal practice. They proposed several models and used structural equation modeling to investigate the relationships of the variables. They were unable, however, to establish adequate fit of the data with the proposed models positioning self-efficacy as either an exogenous or endogenous variable. Exogenous variables are variables that are not influenced by other variables in the model, while endogenous variables depend on other variables and can be both a consequence and antecedent of other variables (Stevens, 2009).

McPherson and McCormick (2006) replicated McCormick and McPherson's (2003) study using a similar group of participants and a similar music performance achievement measure. Participants ($N = 332$) in the 2003 study were ages 9-18 years

while participants ($N = 686$) in the 2006 study were 9-19 years. The Trinity College examination was used as the achievement measure in the 2003 study and the Australian Music Examinations Board examination as the achievement measure in the 2006 study. McPherson & McCormick (2006) constructed a model in which self-efficacy mediated the influence of the variables grade level, practice regulation, formal practice, and informal practice on music performance achievement. They found their data had adequate fit with this proposed model, χ^2 (364, $N = 686$) = 1837.78, $p < .01$, GFI = .94, AGFI = .93, and RMSEA = .08. This finding suggested that self-efficacy influenced other variables commonly associated with music performance achievement.

Studies have found inconsistencies in the interpretations of self-perceptions among diverse groups of individuals. Hendricks (2009) discovered that music students with high self-efficacy beliefs interpreted information related to the sources of self-efficacy differently than students with low self-efficacy beliefs. She found that students with high self-efficacy beliefs reacted strongly to information related to mastery experiences while students with low self-efficacy beliefs were more sensitive to vicarious experiences. She added that high self-efficacy students were positively influenced by positive and negative feedback, peer encouragement, witnessing other students succeed, and issues related to fatigue. In contrast, she found the confidence of students with low self-efficacy beliefs improved after watching other students struggle. These findings were consistent with Austin and Vispoel's (1998) conclusion that "individuals do not attribute success and failure to the same causal factors" (p. 40). They found that students with higher music self-concept attributed their success to ability and family influences, while students with lower music self-concept attributed their failure to lack of ability and lack

of persistence. These findings have instructional implications for music educators. They suggest that the one-size-fits-all approach is not effective. Educators need to differentiate their instruction between students with high and low self-perceptions.

As far as I have found, there has been little investigation into the relationships among aptitude, self-efficacy, and performance achievement in music. In a study of writing achievement however, Pajares, Miller, and Johnson (1999) used multiple regression to determine that writing aptitude ($\beta = .40$, $p < .05$) and writing self-efficacy ($\beta = .39$, $p < .05$) made significant contributions in predicting writing achievement. They also found that writing self-efficacy mediated the relationship between writing aptitude and writing achievement. There is no information on whether this relationship applies in music, but the study raises questions that are worthy of investigation. An investigation of these constructs in music would fill gaps of fundamental knowledge in music aptitude, music self-efficacy, and music achievement.

Summary

Bandura's social cognitive theory provides a framework for investigating self-perceptions. He identified self-efficacy as a primary determinant of human behavior and defined it as an individual's judgments of his or her abilities to accomplish specific tasks. Educational researchers in a variety of subjects such as math, science, and writing have constructed self-efficacy scales based on Bandura's framework. The data generated by these scales have adequately fit Bandura's proposed model, suggesting that his theory provides an accurate representation of cognitive and behavioral processes. These investigations have contributed to improving the understandings of relationships between constructs such as self-efficacy and achievement that are central to the learning process.

Studies in music education have identified the need to investigate self-perceptions in music. Greenberg (1970), Svengalis (1978), and Schmitt (1979) presented evidence showing the influence of self-perceptions on student achievement and attitudes in music. McPherson and McCormick (2000) and McCormick and McPherson (2003) adopted Bandura's framework and demonstrated that self-efficacy was the best predictor of achievement among several constructs. They also found (McPherson & McCormick, 2006) that self-efficacy mediated the influence of several variables such as formal music practice that were fundamental to music learning. In these studies, the researchers examined self-efficacy as a construct held by individuals prior to their participation in the study. This study extended these findings by investigating the sources of information that contribute to the development of self-efficacy beliefs.

I reviewed several instruments designed to measure self-efficacy in music performance and only found one instrument to measure the sources of self-efficacy among band, chorus, or string orchestra students. Other researchers have created scales to measure self-efficacy in jazz (Wehr-Flowers, 2007), for use before a performance evaluation (Hendricks, 2009), and to evaluate general music self-efficacy (Ritchie & Williamon, 2007). The only instrument I found to measure the sources of self-efficacy was the Music Performance Self-Efficacy Scale (Zelenak, in press). This scale had been evaluated with a homogenous group of participants, but needed to be examined with a more diverse group of participants. Therefore, this study met the needs of music education on two levels: (a) by developing new understandings of self-efficacy in music performance, and (b) evaluating the Music Performance Self-Efficacy Scale.

Chapter 3: Method

Research Design

This study used a descriptive design to investigate self-efficacy in music performance and to establish evidence with which to determine the validity and reliability of the Music Performance Self-Efficacy Scale (MPSES). The descriptive information derived from this study was then used to construct new understandings of self-efficacy in music performance. Previous studies examined the relationship between self-efficacy and music performance achievement (McPherson & McCormick, 2000) and self-efficacy's role as mediator between various constructs and music performance achievement (McPherson & McCormick, 2006). This study, however, investigated the sources of information that contribute to the development of self-efficacy beliefs and the capability of the MPSES to measure those sources.

Initially, the investigation centered on the analysis of data generated by the MPSES. I examined the distribution of scores, the characteristics of the scores produced by each item, the relationships between the items and the sources of self-efficacy, and the relationship between the sources of self-efficacy and the composite construct of self-efficacy. The data were evaluated to determine whether the MPSES produced data that accurately represented Bandura's theoretical model. Mean scores were then compared among the independent variables of ensemble and grade level to identify any differences. In addition, scores from the MPSES were tested to establish if the scale items were invariant (i.e., unbiased) in measuring the sources of self-efficacy perceptions across

middle and high school participants. The methods and techniques used in this study have been used in previous analyses of self-efficacy scales (Lent, Lopez, & Bieschke, 1991; Matsui, Matsui, & Ohnishi, 1990; Usher & Pajares, 2006, 2009). In addition, I examined the relationship between music aptitude and self-efficacy in music performance.

This study provided evidence to determine the validity and reliability of the MPSES. To establish evidence of test content, items were examined by a panel of experts in music education. To establish evidence based on response processes, music teachers provided self-efficacy ratings for each participant and these ratings were correlated with the participant's total score on the MPSES. To establish evidence based on internal structure, results of the MPSES were analyzed to determine if they fit Bandura's proposed model. To establish evidence based on relations to other variables, scores from the MPSES were correlated with scores from Schmitt's (1979) Self-Esteem of Music Ability (SEMA) to establish convergent evidence of validity. MPSES scores were also correlated with scores from Usher and Pajares's (2009) Sources of Middle School Mathematics Self-Efficacy Scale (SMSES) to establish discriminant evidence of validity. No evidence was collected regarding the consequences of testing, but teachers were advised not to use results from the MPSES as an achievement measure or to share individual results with students. Results were intended to identify student strengths and weaknesses in order to drive instructional choices. To establish evidence of reliability, I conducted a test-retest evaluation and examined the internal consistency of the item responses within each source of self-efficacy, and within the total scale.

Population and Sample

After receiving approvals from music teachers, school principals, school districts, and the University of South Florida's Institutional Review Board, packages containing the research documents were delivered to the schools. The packages contained instructions to administer the study, the introductory script, parent permission forms, assent to participate forms (for students), an informed consent form (for the teacher), and self-efficacy rating forms. Some packages were delivered by the researcher to schools that were within reasonable driving distances (< 50 miles) while others were mailed to schools farther away. Prior to sending the packages, each teacher was asked to provide the researcher with the number of potential participants they would like to include in the study. This number was used to determine the total number of parent permission forms and student assent forms to send. One thousand five hundred and thirty-nine parent permission forms were sent out. Based on an estimated return rate of 50% for the parent permission forms, 770 assent to participate forms were also sent.

Sixteen music teachers at 15 schools agreed to participate in the study. Packages were delivered to all teachers. Online questionnaire responses, however, were only received from students in 10 schools. Multiple attempts were made to contact missing teachers by email and phone. Three hundred and forty-nine responses were collected. Of those responses, 11 were tests conducted by the researcher and music teachers to check the functionality of the web site, 18 contained only demographic information, and 27 were problematic. Of those problematic responses, some participants accessed the web site on multiple occasions even though the web site offered an option to save the participant's work and provided an email link to resume from the point they left off. By

matching demographic information, I was able to identify and delete those participants with multiple entries. Some submitted partial responses the first time, then submitted complete responses the second. In those cases, I deleted the partial responses. Some completed the entire battery of questionnaires on multiple occasions. In those situations, I kept the first entry and deleted the later entries. And in one case, the participant responded with scores of 100 on all items of the MPSES and SMSMSES—including items on the SMSMSES that were reverse coded. This participant's responses were suspicious and deleted. Finally, three outliers were identified as having *z* scores > $\pm$ 3 and deleted. Two hundred ninety responses were included in the study with an overall response rate of 19%.

Table 5 provides information regarding the response rates per school. The percentage per school was calculated by dividing the number of responses from each school by the number of potential participants offered by the teachers. For example, if a teacher offered to conduct the study with 50 students, and 10 students provided responses, the percentage for that school would be 20%. This number provided insight into the level of involvement within each school. The following column, % of total participants, is the number of participants per school divided by the total number of participants ($N = 290$). This number provided information about the level of involvement between schools. School 1 had the largest number of participants and consequently the greatest percentage of total participants. This was the only school with two music teachers participating in the study. School 8 had the greatest percentage of participation within the school. This teacher offered the study to 46 students and 31 responded.

Table 5
Response Rates by School (N = 290)

Region/School	Grade[a]	Ensemble(s)	*n*	% per school[b]	% of total participants[c]
Southeast					
School 1[d]	HS	Band/Chorus/Orch	98	37	33.79
School 2	HS	Band/Orch	8	8	2.76
School 3	HS	Chorus	3	5	1.03
School 4	MS	Chorus	42	31	14.48
School 5	MS	Chorus	48	38	16.55
School 6	MS	Band/Orch	17	24	5.86
School 7	MS	Orch	11	17	3.79
West					
School 8	HS	Band	31	67	10.69
School 9	MS	Chorus	26	11	8.97
School 10	MS	Chorus	6	7	2.08
Totals	4 HS 6 MS	4 Band 6 Chorus 4 Orch	290	19[e]	100.00

[a] MS (i.e., middle school) consists of students in grades 6-8 (ages 11-14), HS (i.e., high school) consists of students in grades 9-12 (ages 14-18).
[b] The number of responses divided by the number of potential participants as determined by each music teacher.
[c] The number of participants from each school divided by the total number of participants ($N = 290$).
[d] Two teachers from this school participated in the study, one band, and one chorus/orchestra.
[e] The total response rate (i.e., the total number of responses divided by the number of permission forms distributed).

Participants ($N = 290$) were secondary school music students in band, chorus and string orchestra ensembles. Schools were located in the west (i.e., Colorado) and southeastern (i.e., Florida) regions of the United States. Schools in the west were recommended by colleagues from universities in that region and schools in the southeast

were identified by the researcher. One hundred and twenty-four teachers were contacted via email in schools representing diverse populations around the United States. Willingness to participate ultimately determined which schools were included in the study. The participants, therefore, constituted a convenience sample. Based on the purposes of this study, an adequate number of participants were obtained from these schools. Crocker and Algina (1986) reiterated Nunnally's (1967) suggestion that an item analysis study should have 5 to 10 times as many participants as items. Since there were 24 items on the MPSES, the recommended number of participants should be between 120 and 240. The number of participants in this study exceeded those recommendations.

Participants had diverse characteristics (Table 6). There were diversities in sex, race, grade level, and music ensemble. A larger number of participants (n = 227) were from the southeast than from the west (n = 63). Females, Whites, middle school students, and chorus students had the greatest number of participants in each demographic category. The largest subgroup of students with common characteristics were middle school white females in chorus from the southeast (n = 76). This subgroup accounted for 26% of the participants. This percentage also implied, however, that 74% of the participants were not members of this subgroup thus providing evidence of diversity among the participants. Although previous research had found differences in self-efficacy based on sex (Pajares & Valiante, 1999), this study refrained from pursuing this line of inquiry to limit the scope of the current study. Differences in music performance self-efficacy between males and females is an important topic that should be addressed in a future study.

Table 6
Demographic Characteristics of Participants (N = 290)

Characteristic	Region		
	Southeast	West	Total
Sex			
Female	169	38	207
Male	58	25	83
Race			
Asian	26	3	29
Black	28	1	29
Hispanic	25	2	27
Native Amer. Indian	1	1	2
Mixed	25	3	28
Other	7	1	8
White	115	52	167
Grade Level			
Middle (6-8)	118	32	150
High (9-12)	109	31	140
Ensemble			
Band	56	31	87
Chorus	121	32	153
String Orchestra	50	0	50
Total	227	63	290

Based on requirements of the Institutional Review Board at the University of South Florida and the school districts, all participants were required to return a signed parent permission form. Failure to meet this requirement excluded an undetermined number of students from the study. Consequently, participants had one common characteristic–the willingness and capability to have a permission form signed by a parent. Gall, Gall, and Borg (2007) acknowledged that children who return parental permission forms differ from children who do not. The rationales behind the participants' willingness and capabilities to return permission forms were beyond the scope of this study. On a separate yet related topic, it was not feasible to randomly select participants

in this study because they were members of nested contexts (i.e., intact music classes). These conditions suggested that the results be approached with caution.

Instruments

Demographic Information (Appendix A). On the first page of the questionnaire web site, participants were asked to provide demographic information. This information included birth date, gender, race, grade level, music ensemble, and music teacher. The birth date was used to match students with their teacher's student self-efficacy ratings. Grade level and music ensemble were used as independent variables to examine differences in self-efficacy scores. Gender, race, and music teacher were used to sort students by region and differentiate between students with the same birth date.

Music Performance Self-Efficacy Scale (MPSES) (Appendix B). This researcher-constructed scale was designed to measure Bandura's (1977, 1986) four sources of self-efficacy in music performance among band, chorus, and string orchestra students. The four sources of self-efficacy were mastery experience, vicarious experience, verbal/social persuasion, and physiological state. The scale had 24 items and was divided into four sections with the items in each section reflecting one source of self-efficacy. Eight items were in the mastery experience section, five items in the vicarious experience section, six items in the verbal/social persuasion section, and five items in the physiological state section. After uploading the items to the web site and pre-testing the response procedure, I realized that grouping the items by source of self-efficacy promoted similar responses for each item within that section. To control the influence of item order, I rearranged the

items so that similar items were separated. All items remained in the scale but were not organized into sections based on their source of self-efficacy.

The items in the current study were developed and tested in a preliminary study (Zelenak, in press). Thirty initial items were based on items found in general self-efficacy tests (Sherer et al., 1982; Pintrich et al., 1991), subject specific tests (Usher & Pajares, 2006; Lent, Lopez, & Bieschke, 1991), and the music education research literature (McCormick & McPherson, 2003; McPherson & McCormick, 2006; Stewart, 2002). Following Bandura's (2006) recommendations, I worded the items to address diversity in generality, strength, and level. Generality referred to variations in breadth of activities and situations. This quality was evident in items such as, "I have had positive experiences performing music in the past" and "I have had positive experiences performing in large ensembles." Strength addressed the degree to which a foundation of prior experiences had been established. This aspect of self-efficacy was assessed by items such as, "I have overcome musical challenges through hard work and practice." Level captured the magnitude or inherent difficulty of the task and was measured by items such as "I have had positive experiences performing simple music" versus "I have had positive experiences performing complicated music." This initial pool of items was presented to a panel of five music education professors and five doctoral students for evaluation to confirm content validity. The items were then sent to another university professor with expertise in this area for further evaluation. The final revision of the Music Performance Self-Efficacy Scale used in the preliminary study (Appendix C) consisted of 24 items: 7 items for mastery experiences (#1-7), 5 items for vicarious experiences (#8-12), 6 items

for verbal and social persuasion (#13-18), 5 items for physiological state (#20-24), and one item to test the participant's accuracy in responding to the items (#19).

In this preliminary study (Zelenak, in press), participants responded to each item by writing a whole number from 0-100 on a line next to each item. This response format had been used among other studies using self-report questionnaires (Usher & Pajares, 2009, Zimmerman & Martinez-Pons, 1990). It differed, however, from Bandura's (2006) recommendation of using intervals of 10 on a 0-100 scale (i.e., 0, 10, 20, 30, ...100). Participants ($N = 293$) were music and non-music students at a middle school in the southeastern United States. I chose not to include a response option of "not applicable" on the questionnaire for three reasons. First, most students in middle school have had some experience performing music either in elementary school or in out-of-school activities. Second, a response option of not applicable would have provided some participants with an escape from thinking deeply about their music self-perceptions. And third, self-efficacy scales in other subject areas had not included a not applicable option. In addition to these reasons, the low self-efficacy ratings from non-music students were essential in determining whether or not the MPSES had the ability to discriminate between music and non-music students.

Results from the preliminary study indicated the MPSES generated scores that represented the four sources of self-efficacy in each section and the composite construct of self-efficacy as an aggregated total. Responses within each of the four sources of self-efficacy were normally distributed. Skewness values ranged from -0.18 to -0.81 and kurtosis values ranged from -0.68 to -1.36. Responses within each section (i.e., source of self-efficacy) were internally consistent: mastery experience $\alpha = .93$, vicarious experience

α = .90, verbal/social persuasion α = .94, and physiological state α = .90. Responses on the total scale were also internally consistent, α = .97. The inter-item correlations ranged from .50 to .70 indicating the items elicited responses that were similar, but not identical. The correlations between the sources of self-efficacy ranged from r = .75 to r = .88, while the correlations between each source of self-efficacy and the composite construct ranged from r = .89 to r = .96 (Table 7). These findings were very encouraging and supported my decision to proceed with the current study.

Table 7
Correlations of Sources of Self-Efficacy Scores and Composite Self-Efficacy Scores from Preliminary Study (N =293)

	Mastery Experience	Vicarious Experience	Verbal/Social Persuasion	Physiological State
Mastery Experience	_			
Vicarious Experience	.77	_		
Verbal/Social Persuasion	.87	.80	_	
Physiological State	.86	.75	.89	_
Composite Self-Efficacy	.95	.88	.96	.94

All correlations were significant (p < .001).

To establish evidence of validity, the results of the MPSES were correlated with results from Pajares' (2007) Writing Self-Efficacy Scale and academic self-efficacy items from Bandura's (1990) Children's Perceived Self-Efficacy Scale. These correlations indicated positive relationships between the construct of music performance self-efficacy

and writing self-efficacy ($r = .30$, $p < .01$) as well as music performance self-efficacy and academic self-efficacy ($r = .34$, $p < .01$). The low magnitude of the correlations suggested the scales were measuring related, yet fundamentally different constructs. This finding was consistent with Bandura's belief that self-efficacy beliefs were domain specific.

Confirmatory factor analysis (CFA) was used to determine whether or not the data generated by the MPSES was consistent with Bandura's proposed model. This procedure was appropriate for evaluating data produced by an instrument based on a preexisting theoretical framework. The results of the initial CFA were χ^2 (226, $N = 293$) = 650.77, $p < .001$, CFI = .93, RMSEA = .08, and SRMR = .04. These values were very close to the benchmarks of good fit recommended by Hu and Bentler (1998) of CFI $\geq$ .95, RMSEA $\leq$ .06, and SRMR $\leq$ .08. In the same article, however, Hu and Bentler mentioned that other researchers advised accepting CFI values as low as .90 (Bentler & Bonett, 1980) and RMSEA values between .05 and .08 (Browne & Cudeck, 1993). Modification indices produced by Mplus identified several sources of misfit. They indicated strong correlations between the error terms associated with items 9 and 10 and items 13 and 20 (Appendix C). These items were also very similar in verbal meaning and content. I allowed the error terms associated with items 9 and 10 and with items 13 and 20 to correlate and the fit indices improved χ^2 (185, $N = 293$) = 475.51, $p < .001$, CFI = .95, RMSEA = .07 and SRMR = .04. Although not meeting all of Hu and Bentler's benchmarks (i.e., a higher amount of variance remained in the error terms as reported in the RMSEA statistic) the results met most values accepted by Hu and Bentler as well as other researchers' recommendations.

To summarize, the preliminary study demonstrated that the MPSES had the ability to produce results that were normally distributed, internally consistent, related to one another, and were related to the composite construct of self-efficacy. In addition, the MPSES was found to measure a unique form of self-efficacy that exhibited modest relationships with academic self-efficacy and writing self-efficacy. Finally, the results adequately fit Bandura's proposed model of self-efficacy from his social cognitive theory.

A few alterations were made to the MPSES in the current study. These alterations were based on recommendations made by an expert panel of 4 music education professors and 7 doctoral students in music education, as well as, results from the preliminary study. Preliminary item 3 (Appendix C) ("I have had positive experiences performing solo, or, in small ensemble) was split into two separate items to reflect the different experiences of performing solo and performing in a small ensemble. Descriptive numbers were also added to promote consistency among the participants' interpretation of the items (i.e., small ensemble 2-10 performers, and large ensemble 11 or more performers). As mentioned above, preliminary item 9 ("I have improved my music performance skills by watching other students who are similar to me in some way, perform well.") was removed due to its similarity to item 10 ("I have used other music students as models to improve my performance skills."). Even though preliminary items 13 and 20 had a strong correlation between error terms, both items were kept in the scale because they reflected different sources of self-efficacy, preliminary item 13 reflected verbal/social persuasion, and preliminary item 20 reflected physiological state. A new item was added to the vicarious experience section to increase the generality of the items

in that section to include the experience of watching someone other than a peer or professional perform music. Preliminary item 19, that was a check of participant response accuracy, was removed from the revised scale. In its place to reduce the effects of item order, the items were rearranged so that similar items reflecting the same source of self-efficacy were no longer next to each other.

In the revised version of the Music Performance Self-Efficacy Scale used in this study (Appendix B), the MPSES consisted of items reflecting the four sources of self-efficacy. Mastery experience consisted of items 1, 4, 6, 8, 10, 12, 14, and 16; vicarious experience consisted of items 2, 5, 11, 18, and 20; verbal/social persuasion consisted of items 3, 7, 9, 13, 21, and 22; and physiological state consisted of items 15, 17, 19, 23, and 24. In addition, the response scale was changed from 0-100 to 1-100. This change was necessitated by a requirement of the web site delivering the questionnaires. One to 100 is an actual 100 point scale, while 0-100 is a 101 point scale.

Sources of Middle School Mathematics Self-Efficacy Scale (SMSMSES) (Appendix D). Usher and Pajares (2009) developed, and evaluated the validity and reliability of their Sources of Middle School Mathematics Self-Efficacy Scale. Although validated among middle school students, I decided to use this scale in the current study that included middle and high school students based on the authors' determination it was psychometrically sound and recommended for adaptation to other domains. The final version of the scale consisted of 24 items that reflected Bandura's four sources of self-efficacy. They used a 3-phase process to construct the scale in which they (a) created the items, (b) conducted an exploratory factor analysis to identify underlying factors and detect problematic items, and (c) subjected the final items to confirmatory factor analysis

to evaluate the fit of the resulting data to Bandura's theoretical model. During the final phase, Usher and Pajares retained only items that met certain requirements. Those requirements were that the items exhibited item-total correlations of $r \geq .55$ with the subscale totals, the inter-item correlation within each source of self-efficacy were $r \geq .30$, and the content of each item was unique. Six items were chosen to reflect each source of self-efficacy and all subsections had internal consistencies of $\alpha > .80$. The data from the final phase demonstrated adequate fit with Bandura's model, χ^2 (246, $N = 803$) = 601.21, $p < .0001$, CFI = .96, RMSEA = .04, SRMR = .04. The standardized factor loadings in the model were significant (i.e., $p < .05$) and they ranged from .61 to .83. The items correlated with the four self-efficacy measures individually and when combined into subscales. Strong relationships between the sources of self-efficacy, motivation constructs, and achievement provided evidence of relationships with other variables. Weak relationships between the sources of self-efficacy and performance approach goal orientation provided evidence that the scale was able to discriminate between different constructs.

The SMSMSES was used to provide evidence of the MPSES's ability to discriminate between the constructs of mathematics self-efficacy and music performance self-efficacy. Both scales were based on the same theoretical model, but each scale measured a different subject domain. Since Bandura suggested that self-efficacy beliefs were tied to specific tasks, I expected a positive, yet weak, correlation between scores on the SMSMSES and the MPSES. The SMSMSES was appropriate for use in this study because it had been rigorously tested in educational settings among participants of age 11-14 years.

Self-Esteem of Musical Ability (SEMA) (Appendix E). Schmitt (1979) developed and validated this scale to measure music self-esteem among children 10-15 years of age. It consisted of 43 items. Participants responded to the items using a 4-point Likert-type scale selecting strongly disagree, disagree, agree, and strongly agree. She suggested the measure was highly reliable based on the correlation of scores from the initial test and a re-test taken three days later (r = .91). Factor analysis indicated the items addressed three separate self-perception factors: (a) self-confidence with a strong influence in music, (b) skills and abilities in music, and (c) students' feelings of acceptance and reinforcement by parents, teachers, and friends (p. 89). These factors have striking similarities to Bandura's sources of self-efficacy. Since self-efficacy and self-esteem are similar constructs, I chose to use the SEMA to establish evidence of convergent validity. Evidence had been established to support the validity and reliability of scores from SEMA (Schmitt, 1979). It has been used in other music studies of self-perceptions (Austin & Vispoel, 1998; Randles, 2010). Due to the similarity of constructs, I anticipated a positive relationship between the scores from each measure.

Although self-esteem of music ability and self-efficacy in music performance are closely related constructs, they have a fundamental difference. Schmitt (1979) described self-esteem of musical ability as "the personal estimation of an individual towards his capacity to perform music tasks and obtain musical competence" (p. 10). Bandura (1986) defined perceived self-efficacy as "people's judgments of their capabilities to organize and execute courses of actions required to attain designated types of performances" (p. 391). In both definitions, the constructs address self-perceptions of skills and abilities, evaluation of those skills and abilities in relation to specific tasks, and the anticipation of

outcomes. The difference between the constructs, however, is in the type of judgment they represent. Bandura (1997) clarified this difference stating "Perceived self-efficacy is concerned with judgments of personal capability, whereas self-esteem is concerned with judgments of self-worth" (p. 11). This difference in judgments separates the two self-perceptions as unique constructs.

Advanced Measures of Music Audiation (AMMA) (Appendix F). Gordon (1989) constructed this instrument to measure stabilized music aptitude. He defined music aptitude as a student's potential to learn music and believed that music aptitude stabilized in individuals after the age of nine. After the age of nine, music instruction had the potential to improve music achievement, but this achievement was limited by one's music aptitude. He defined music achievement as evidence of a person's musical accomplishments. Gordon believed that a student with higher music aptitude had the potential to reach higher levels of music achievement than a student with lower aptitude.

Fullen (1993) found evidence indicating a relationship between music aptitude and music achievement. I chose to include the AMMA in this study because of this relationship between aptitude and achievement and the potential influence of self-efficacy on music performance achievement. Although originally intended for use among undergraduate music students, the AMMA has also been found to provide accurate results with middle school students (Gordon, 2004) and high school students (Fullen, 1993). Its use was appropriate for the age of the participants in this study and can be administered to a variety of music students regardless of music instrument, music preference, or level of experience.

The AMMA is a listening test. It has 30 items with each item consisting of two musical phrases. The participant listens to the two phrases and decides whether the phrases are the same or different. If a phrase is different, the participant must determine whether the source of the discrepancy is tonal or rhythmic. In a study using a pre- and post-test design, Fullen (1993) found no significant difference in AMMA scores among junior and senior high students following three weeks of music instruction. This finding supported Gordon's belief that music aptitude was stabile in individuals over the age of nine. Fullen also found the correlations between initial AMMA scores and scores collected ten days later were moderately strong for students in grades 7-8 (r = .72) and somewhat stronger for students in grades 9-12 (r = .81). These findings verified the stability of music aptitude over time. In addition, Fullen uncovered positive correlations between AMMA scores and vocal performance test results for students in grades 7-8 (r = .25) and in grades 9-12 (r = .24). These findings suggested that the AMMA was a measure of music aptitude that had predictive validity in determining music performance achievement. The validity and reliability of the scores from this measure were another factor that influenced my decision to include it in the current study. AMMA scores were used as the predictor variable in a regression equation to determine the influence of music aptitude on self-efficacy in music performance.

Teacher ratings. Music teachers rated their students' music performance self-efficacy on a 1-5 scale (1 = low, 5 = high). They were asked to draw on their own perceptions of the students' behaviors in music performance contexts and indicate the relative strengths of the students' beliefs in their abilities to perform music. For example, a student who volunteers frequently to perform would be perceived as having a stronger

belief in their ability to perform than a student who does not volunteer. Participants were instructed to put their name and birthdate on the teacher's Self-Efficacy Rating Form (Appendix G). Once the teacher had completed their ratings, they would remove the participants' names from the form to ensure confidentiality and return the form to the researcher. I used the birthdates to match the scores on the Self-Efficacy Rating Form with student scores on the online questionnaires.

Data Collection

The data collection process followed procedures recommended and approved by the University of South Florida's Institutional Review Board and the individual school districts. Music teachers were sent a package that included: Instructions for Self-Efficacy in Music Performance Study (Appendix H), Introduction Script (Appendix I), Self-Efficacy Rating Forms (Appendix G), Parent Permission Forms (Appendix J), Assent to Participate in Research (Appendix K), and an Informed Consent to Participate Form (Appendix L). Upon receiving the package, teachers were instructed to read the Introduction Script to their students. The script introduced the researcher, the objectives of the study, and the responsibilities of the participants. Students were then given Parent Permission Forms to take home. Those students returning signed Parent Permission Forms were then advised to read and sign the Assent to Participate in Research and to add their name and birthdate to the Self-Efficacy Rating Form. After completing this process the students were eligible to participate in the study.

Data were collected using an Internet-based survey company (SurveyGizmo.com). The company provided a web site on which the participants responded to the questionnaires, allocated server space for storing audio files, and

facilitated the export of raw data. The web site was initially a blank template that I redesigned to meet the needs of this study. Data were collected in the following order: Demographic Information (Appendix A), Music Performance Self-Efficacy Scale (Appendix B), Sources of Middle School Mathematics Self-Efficacy Scale (Appendix D), Self-Esteem of Musical Ability (Appendix E), and Advanced Measures of Music Audiation (Appendix F). Participants responded to items on the Music Performance Self-Efficacy Scale and the Sources of Middle School Mathematics Self-Efficacy Scale by entering a number from 1-100 in the answer box. They responded to items on the Self-Esteem of Musical Ability by clicking on their selection of strongly disagree, disagree, agree, or strongly agree. They clicked on the corresponding box of "Same," "Tonal," or "Rhythm" for the Advanced Measures of Music Audiation. Participants listened to audio files, determined whether the pair of musical phrases was the same or different, and then made their response selection. The study questionnaires were presented to all participants in the same order. Since items on the MPSES an SEMA were similar, it was necessary to separate the two questionnaires to control test fatigue. Also, the Advanced Measures of Audiation required the most time to complete and was therefore placed at the end of the questionnaire battery. Participants were given the option to save their partial responses and return to the questionnaire if they did not have time to complete the entire battery. Another rationale for placing the AMMA at the end of the data collection process was to maintain the concurrent nature of the responses from the three self-perception questionnaires. Since music aptitude had been shown to be stable (Fullen, 1993; Gordon, 1989), returning to that questionnaire at a later time would have minimal influence on the results.

Data Analysis

Questionnaire results were downloaded from the web site as Excel comma-separated value files (i.e., .cvs). This format could be read by the SPSS and SAS software programs. Data to be read by the Mplus software were loaded into Excel and converted to a text file (i.e., txt). Descriptive information such as means, standard deviations, and frequency distributions were calculated for all measures. Additional analyses were conducted on the results from the MPSES. These analyses included inter-item correlations, item-factor correlations, item-total correlation, and internal consistency. Prior to addressing the research questions, results were examined for outliers, missing values, and normal distribution. In addition, the data were examined to determine whether they met the underlying assumptions for the analytical procedures.

Research Question 1: Do the responses of secondary school music students to items on the Music Performance Self-Efficacy Scale adequately fit Bandura's theoretical model of self-efficacy?

Confirmatory factor analysis was used to determine whether the data resulting from the Music Performance Self-Efficacy Scale fit Bandura's theoretical model. There were 1st and 2nd order levels in this model (Figure 3). The 1st order consisted of the questionnaire items (i.e., variables), the sources of self-efficacy (i.e., factors), their error terms, and the loadings between each of these elements. The 2nd order was an extension of the first and included the composite construct of self-efficacy, the sources of self-efficacy, their error terms, and the loadings between each of these elements. Mplus version 6 software was used to determine values and relationships. Hu and Bentler's

(1998) recommendations were used as benchmarks to determine the goodness of fit, CFI $\geq$ .95, RMSEA $\leq$ .06 and SRMSR $\leq$ .08.

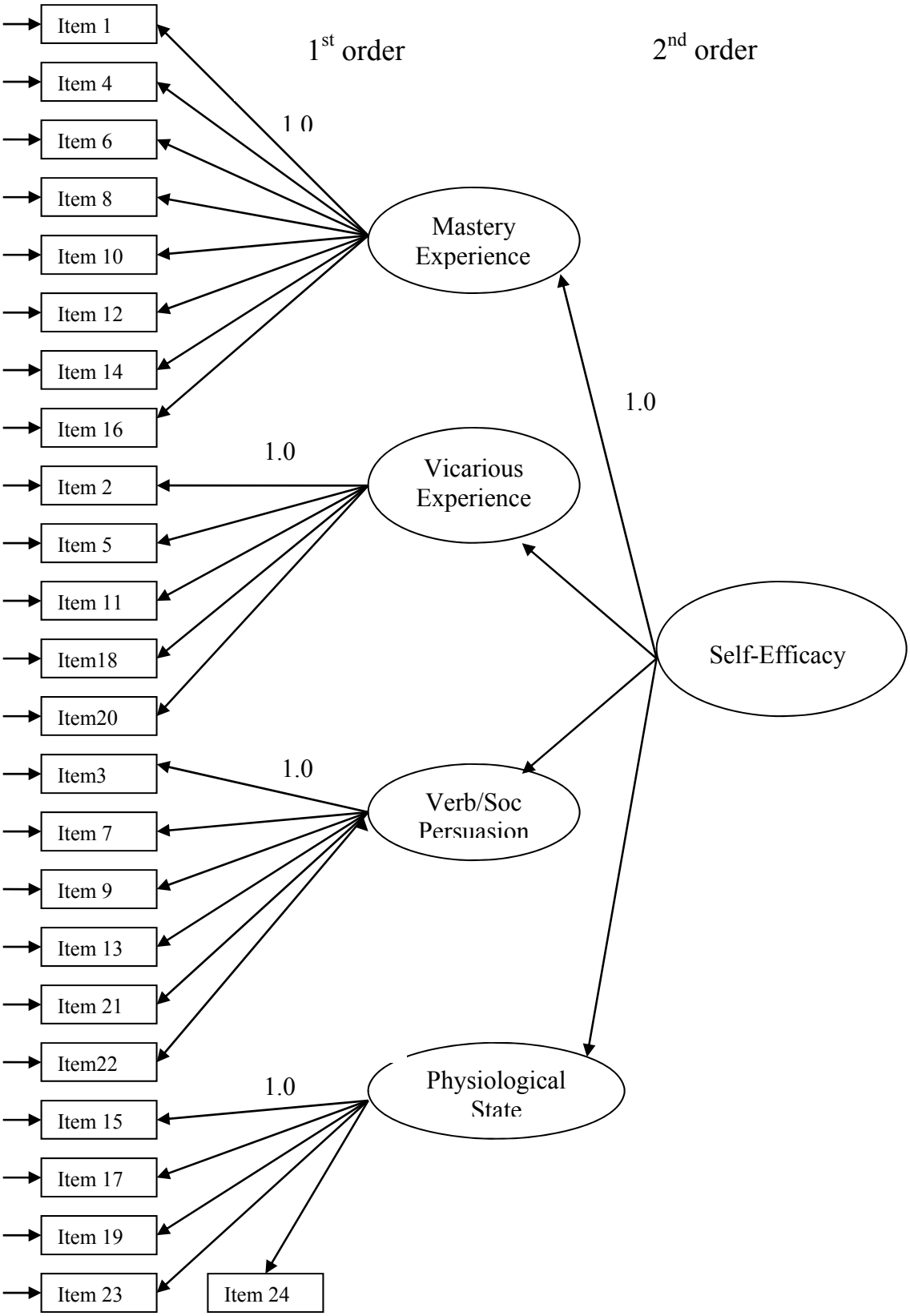

Figure 3. Model Identification of the Music Performance Self-Efficacy Scale

The first step of the CFA was to specify the model. The specification of the model was based on Bandura's framework as outlined in his social cognitive theory with each item on the questionnaire acting as a reflective indicator of one of four sources of self-efficacy and those four sources contributed to the composite latent construct of self-efficacy.

The second step was to identify the model. Since there were 24 items on the questionnaire, there were 300 values in the covariance matrix. The loading from one item for each factor was set to the value of 1.0. Those loadings were item #1 (mastery experience), item #2 (vicarious experience), item #3 (verbal/social persuasion), and item #15 (physiological state). Setting these loadings to 1.0 allowed the software to estimate the other loadings of the model. Similarly, in the 2nd order model, the loading from mastery experiences to self-efficacy was set to 1.0, allowing these loadings to be estimated as well.

The third step was to estimate the model. In this study, maximum likelihood estimation was used to estimate the free parameters. As an outcome of the identification process in step two, I estimated 52 parameters resulting in 248 degrees of freedom in the baseline 1st order model. The 52 estimated parameters included the loadings between the 24 items and the four 1st order latent variables, the errors associated with the items, the loadings between the latent variables and the 2nd order composite construct, and the errors associated with the latent variables, the error associated with the latent construct. Mplus software generated the standardized loadings between the factors and items in the fifth column of the output labeled "StdYX". These standardized loadings were similar to

standardized regression coefficients with the advantage that they can be compared to one another (Thompson, 2004).

The fourth step was to use Mplus to produce fit indices. Since the sample size was large ($N = 290$), the χ^2 statistic was sensitive to sample size and did not accurately represent the fit of the data. The comparative fit index (CFI), the root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMR) exhibited greater accuracy with larger sample sizes (Hu & Bentler, 1998). The CFI indicated the degree of similarity between the variance and covariance matrices of the data set and the proposed model, while the RMSEA and SRMR indicated the lack of fit between the error variances of the data and the proposed model. Therefore, larger CFI values were anticipated indicating more similarity between the data and the model, while smaller RMSEA and SRMR values were anticipated reflecting fewer differences.

After examining the fit of the model, the modification indices (MI) were reviewed. Items with high modification index values indicated optimum places to make changes to the model. Adding correlations between the error terms associated with the items, or correlations between the factors are a couple of ways of reducing variance in the model and improving the fit. If after three modifications, the fit is not within the commonly accepted ranges for good fit mentioned above, there may be a more serious problem with the items, the sample size, or the overall model. Fit and index values are helpful in determining whether or not modifications must be made to the model.

Research Question 2: How much influence do mastery experience, vicarious experience, verbal/social persuasion, and physiological state have on self-efficacy in music performance among secondary school music students?

Confirmatory factor analysis (CFA) was used to determine the factor loadings. In simplest terms, loadings are like regression coefficients. They represent the percentage of change in one element based on a unit of change in another. First order loadings signified the strength of the relationship between each item on the questionnaire (i.e., variable) and each source of self-efficacy (i.e., factor). Second order loadings characterized the relationship between each source of self-efficacy and the composite construct. The 2nd order loadings were then compared to conclude whether or not the influence of each self-efficacy source was significant and to determine the relative degree of influence attributed to each source. To avoid capitalizing on chance, Stevens (2009) recommended testing each loading for significance at $\alpha = .01$ and he provided a chart of critical values (p. 332) for use in the analyses.

Research Question 3: Are the items on the Music Performance Self-Efficacy Scale invariant for middle and high school students?

Once good fit was established between the data from the MPSES and Bandura's proposed model, I conducted factor invariance testing on the 24 items of the MPSES across middle and high school participants. The intent of this test was to confirm that the scale items were not biased toward one grade level or another while maintaining the assumption that all participants had equal levels of ability. There were several steps in conducting a factor invariance test. First, I established baseline data for the model. Separate confirmatory factor analyses (CFA) produced χ^2 and degrees of freedom values for middle and high school participants. The χ^2 values and the degrees of freedom from each group were combined to create the baseline values. In the following steps, various parameters of the model were constrained (i.e., held constant) and the resulting χ^2 and *df*

values were compared either to the baseline values, or the values of the preceding step depending on the location of the step in the process. If the difference in values were significant, then there was a difference in the functioning of the items for the two types of ensembles. For example, if more high school participants consistently rated an item higher than middle school participants, there may be some characteristic of the item that measures the two groups of participants differently. When there was no significant difference, the comparison was considered tenable and the data were invariant at that step of analysis. Since the data was determined to have a non-normal multivariate distribution, the Satorra-Bentler scaled chi-square difference was used to represent the change in χ^2 between models (Dimitrov, 2010). This value does not represent the change in Satorra-Bentler χ^2 values ($SB\Delta\chi^2$) which is not distributed as χ^2, but makes adjustments to these values in order to follow the χ^2 distribution.

In the second step, the factor loadings were constrained across groups. The resulting degrees of freedom and χ^2 were compared to the baseline value. This procedure established the metric invariance of the model and whether the items related to the constructs in the same way across the two groups. When the difference was not significant, the metric invariance was tenable between the sets of data.

The third step tested the scalar invariance between the models. In this step, the intercepts (i.e., the mean scores for each group) were constrained to be equal and the change in SBS$\Delta\chi^2$ and degrees of freedom indicated whether or not this constraint was tenable. The purpose of this step was to determine if the questionnaire measured the same level of difficulty for both groups. If there was a significant difference, then constraining

the intercepts to be equal was not tenable and there was a difference in the endorsement of the questionnaire between the two groups.

The following steps were a continuation of the previous steps with each step progressively constraining another parameter in the model. In the fourth step, the error variances were constrained to be equal. In the fifth step, the covariance of the latent variables was constrained to be equal. And finally in the sixth step, the variance of the latent variables was constrained to be equal. When the models made it through each of these steps without indicating a significant difference, then the assumption that the questionnaire was invariant and operating identically across the groups was tenable..

Research Question 4: Are there differences in the value attributed to mastery experience, vicarious experience, verbal/social persuasion, and physiological state by band, chorus, and string orchestra students?

Multivariate analysis of variance (MANOVA) was used to compare the composite self-efficacy scores among the types of ensembles (i.e., band, chorus, and string orchestra). That analysis was followed by univariate analysis of variance (ANOVA) using each source of self-efficacy as separate dependent variables. The Tukey pairwise comparison test was then used to determine if there were significant differences between types of ensembles.

Research Question 5: Are there differences in the value attributed to mastery experience, vicarious experience, verbal/social persuasion, and physiological state by middle and high school music students?

Once again, multivariate analysis of variance (MANOVA) was used to compare the different levels of the independent variable on the composite self-efficacy scores.

Univariate follow-up test were then used to compare the mean scores of middle and high school participants on each source of self-efficacy.

Research Question 6: To what extent does music aptitude predict self-efficacy in music performance?

Scores from the AMMA and MPSES were entered into a regression equation. Scores from the AMMA were used as the predictor variable and scores from the MPSES as the criterion variable. This equation produced a coefficient that signified the relationship between music aptitude and self-efficacy in music performance. The regression equation also produced R^2 information that indicated the percentage of variance that could be predicted in self-efficacy based on music aptitude.

Research Question 7: To what extent does evidence support the use of the Music Performance Self-Efficacy Scale as a measure of self-efficacy in music performance?

To establish evidence based on test content, the items were examined by a panel of experts in music education and self-efficacy prior to conducting the study. This panel consisted of four music education professors and seven doctoral students in music education. To establish evidence based on response process, scores from the MPSES were correlated with the teachers' ratings of student self-efficacy in music performance. These ratings were intended to reflect the student's willingness to perform in class and in other situations demonstrating a high sense of self-efficacy for music performance. To establish evidence of internal structure, confirmatory factor analysis was used to determine whether the data generated by the MPSES adequately fit Bandura's proposed model. To establish evidence base on relations to other variables, scores from the MPSES were correlated with scores from Schmitt's (1979) Self-Esteem of Music Ability. Since

the latent variables were similar, this correlation provided a source of convergent evidence. Another source of convergent evidence was in the correlation between MPSES scores and teacher ratings of student self-efficacy. Scores from the MPSES were also correlated with scores from Usher and Pajares's (2009) Sources of Middle School Mathematics Self-Efficacy Scale. Since these scales measure self-efficacy in different subject areas and self-efficacy was proposed to be subject specific (Bandura, 1997), this relationship demonstrated evidence of discriminant validity. To establish evidence of reliability, the internal consistency of the item responses were examined for each item, each section of the MPSES, and the total scale.

Software.

The computer software Statistical Analysis System 9.2 (SAS), Statistics Program for the Social Sciences 19 (SPSS), and Mplus version 6 were used to analyze the data.

Chapter 4: Results

This chapter is organized in four sections (a) data characteristics, (b) item evaluation of the Music Performance Self-Efficacy Scale, (c) analytical assumptions, and (d) research question results. The data characteristics section contains general information about the data such as missing data, outliers, and reliability estimates. The item evaluation section is a detailed examination of the responses generated by the MPSES. It begins with a discussion of the individual items, followed by the sources of self-efficacy, and then the composite construct of self-efficacy in music performance. The analytical assumptions section provides information regarding the assumptions underlying the analytical techniques used to answer the research questions. These assumptions include independence, normal distributions, homogeneity of variance, and homogeneity of the covariance matrices. Finally, in the research question results section, I use the findings of the study to answer the research questions. Each research question is presented, a brief description of the procedure is provided, results are given, and an interpretation of the results follows.

Data Characteristics

Missing data.

Two hundred and ninety participants completed the Music Performance Self-Efficacy **S**cale. Since it was the first questionnaire in the battery, there were no problems with missing data. Participants were not allowed to move on to the next questionnaire

until all of the items on the current questionnaire had been completed. This feature was offered by the web site provider.

The battery of questionnaires was lengthy requiring 30-60 minutes. Some participants were not able to complete all questionnaires during their class periods. All participants completed the MPSES (N =290) and most completed the Sources of Middle School Mathematics Self-Efficacy Scale (n = 289), the Self-Esteem of Musical Ability (n = 286). The Advanced Measures of Music Audiation was the last questionnaire in the battery and fewer participants completed it (n = 222). It required listening to audio examples and ran approximately 20 minutes. It was intentionally placed as the last questionnaire to reduce missing data in the other questionnaires. If it had been placed second or third, there was a greater possibility of missing data in the SMSMSES and SEMA. In the analysis using data from the AMMA, only participants with complete response sets were included.

Outliers.

The responses from three participants were identified as outliers on the MPSES. These responses were considered outliers because their z scores were at the extreme ends of the normal distribution (z scores: -4.60, -3.93, and -3.83). A z score represents the number of standard deviations from the mean. Based on a normal distribution, most scores (approximately 96%) fall within two standard deviations above or below the mean. Although these outliers could occur in the larger population, the probability was small. In addition, these outliers had a powerful influence on the distribution of the MPSES scores. With the three outliers included, the distribution was not normal. When removed, the distribution fell within acceptable limits of normal distribution (skewness +/- 1, kurtosis

+/- 2). I analyzed the distributions with and without the outliers using skewness and kurtosis as indicators of normality. Theoretically, the closer the value of each indicator was to zero, the more normal the distribution. Skewness with the outliers included was -1.16 and without the outliers was -0.70. Kurtosis with the outliers included was 2.18 and without the outliers was 0.13. Stevens (2009) advised removing outliers stating, "Statistical procedures in general can be quite sensitive to outliers" (p. 11). The purpose of this study was not to evaluate participants with extreme self-perceptions of self-efficacy, but rather to examine the self-efficacy beliefs of groups of participants. I removed the outliers from the analysis. Two hundred ninety participants remained after removing the outliers.

Reliability.

One form of reliability refers to an instrument's ability to replicate results when taken by the same participant on multiple occasions under similar conditions. This form of reliability can be established by comparing the results from alternate test forms or through the test-retest method. The test-retest method consists of administering the same test to the same participants on separate occasions under similar conditions. The correlations between the participants' test results are then used to determine the "coefficient of stability" (Crocker & Algina, 1986, p. 133). This coefficient provides information about the instrument's ability to replicate scores and the stability of the construct.

The test-retest procedure was used with a small group of middle school band and string orchestra participants (n = 14). Since the current study was conducted at the end of the school year, it was not possible to increase the number of participants due to

performance obligations, limited access to computer labs, and lack of time. These participants completed the MPSES on two separate occasions approximately 3 weeks apart. The correlation of results from the first and second administrations of the MPSES was $r = .87$. The strength of this relationship was encouraging even though the number of participants was small and there were many distractions in the school. The correlation between the test and retest results for the sources of self-efficacy were mastery experience $r = .91$, vicarious experience $r = .71$, verbal/social persuasion $r = .78$, and physiological state $r = .60$. The rank order of correlations was similar to the rank order of the number of questions reflecting each source. Mastery experience demonstrated the strongest correlation and had eight items, verbal/social persuasion exhibited the second strongest correlation and had six items, followed by vicarious experience and physiological state with five items each. The small number of participants ($n = 14$) may have also influenced the strengths of the correlations.

Internal consistency is another form of reliability. It is an estimate of item homogeneity, or "the degree to which items on a test jointly measure the same construct" (Henson, 2001, p. 177). Cronbach's coefficient alpha is a statistic commonly used to report internal consistency. The closer alpha is to 1.0, the greater the consistency among the scores. The internal consistencies of the measurement instruments used in this study are reported in Table 8. Among the sources of self-efficacy on the MPSES, responses reflecting verbal/social persuasion were the most consistent ($\alpha = .77$), while responses to items reflecting vicarious experiences were least consistent ($\alpha = .59$). The internal consistency for all responses on the MPSES was $\alpha = .88$. It is not unusual for responses from the sections of a scale to exhibit less internal consistency than the consistency of the

entire scale. Only the score totals from the SMSMSES, SEMA, and the AMMA were used in the analyses. Therefore, only the internal consistencies for the score totals were included in Table 8.

Table 8
Internal Consistency Estimates (N = 290)

Source	α	95% CI[a]
Mastery Experience	.74	.69 to .78
Vicarious Experience	.59	.51 to .66
Verbal/Social Persuasion	.77	.73 to .81
Physiological State	.67	.61 to .73
MPSES total	.88	.86 to .90
Mathematics Self-Efficacy (n = 289)	.81	.77 to .84
Self-Esteem of Music (n = 286)	.95	.94 to .96
AMMA (n = 222)	.89	.87 to .91

[a] CI = confidence interval.
AMMA = Advanced Measures of Music Audiation.

The internal consistency of responses from all scales used in this study exceeded the benchmark of α =.80 for general research purposes (Henson, 2001; Nunnally, 1967). A comparison of the alpha values among scales indicated that the internal consistency of the MPSES was greater than that of the SMSMSES (α = .81), but not as great as the AMMA (α = .89). Responses on the Self-Esteem of Musical Ability, however, were the most consistent (α = .95) of all scales.

Item Evaluation of the Music Performance Self-Efficacy Scale

Item statistics were grouped according to source of self-efficacy in Table 9. Participants rated themselves from 1-100 on each of the 24 items of the MPSES (Appendix B). Item 8 had the highest mean score (M = 89.19, SD = 16.98). Many

participants reported having positive experiences performing simple music. In contrast, item 19 had the lowest mean score (M = 58.32, SD = 34.59) indicating many school aged music students worried about making small mistakes during their performances. Item 19 was the only item having any form of negative syntax among the MPSES items. There may have been confusion among participants in how to respond to this item. The 30 point gap between the highest and lowest mean scores indicated that the items on the MPSES covered a variety of self-perceptions and allowed for a wide range of responses.

Table 9
MPSES Item Statistics Grouped by Source of Self-Efficacy (N = 290)

Source Item	Item-Source Correlation[a]	Item-Total Correlation[b]	*M*	*SD*
Mastery Experience				
1. ...positive experiences performing music in the past	.55	.53	86.51	18.66
4. ...positive experiences performing in large ensembles.	.37	.39	87.73	17.71
6. ...positive experiences performing solo.	.49	.55	59.53	35.96
8. ...positive experiences performing simple music.	.31	.33	89.13	16.98
10. ...positive experiences performing complicated music.	.53	.54	78.59	22.28
12. ...overcome musical challenges through hard work and practice.	.49	.55	80.97	22.97
14. ...used a practice routine to prepare for my performance.	.37	.44	60.49	32.86
16. ...positive experiences performing in a small ensemble.	.50	.53	71.93	31.77
Vicarious Experience				
2. ...improved skills by watching professional musicians...	.35	.43	68.26	28.74
5. ...improved skills by watching someone who is similar to me...	.35	.42	67.97	32.56
11. ...used other students as models	.38	.29	71.50	30.39

18. ...watched other students ...and then decided whether I could perform the piece...	.32	.33	74.76	30.19
20. ...compared my performance skills with those of other students	.33	.30	78.91	26.45
Verbal/Social Persuasion				
3. My friends think I am a good performer...	.59	.60	79.33	23.90
7. Members of my family believe I perform well.	.51	.49	86.60	21.92
9. People have told me my practice efforts have improved my performance skills.	.45	.54	77.40	27.97
13. ...received positive feedback on music performance assessments.	.50	.48	85.65	17.96
21. My music teacher has complimented me on my musical performance.	.55	.55	78.61	27.71
22. I have met or exceeded other people's expectations of being a good musician...	.56	.53	73.04	27.97
Physiological State				
15. I am learning to control nervousness...	.39	.34	78.26	26.91
17. Performing with my instrument/voice makes me feel good.	.51	.62	84.36	25.50
19. I do not worry about small mistakes during a performance.	.26	.25	58.32	34.59
23. I enjoy participating in musical performances.	.56	.62	88.22	22.77
24. I have positive memories...of my past musical performances.	.52	.53	85.81	22.19

a. Correlation with the source of self-efficacy score (i.e., mastery experience, vicarious experience, etc.).
b. Correlation with composite self-efficacy score.

Differences were also found in the amount of variance associated with each item. These differences indicated that not all participants held the same self-perceptions. Item 6 elicited the widest range of responses resulting in a larger amount of variance than the other items (*SD* = 35.96). Participants held a wide range of self-perceptions in regards to their experiences performing solo. This may have been a reflection of the music classes in which they were enrolled. Band, chorus, and string orchestra classes focus primarily on large ensemble performance. The participants may not have had instruction or opportunities to perform solo in these classes. One other explanation may be the developmental stage at which these participants were in. Adolescents often avoid situations in which they can be singled out in order to avoid criticism from their peers. In contrast, responses to item 8 were the least diverse (*SD* = 16.98). Most participants rated their experiences of performing simple music highly.

Aside from the individual characteristics of the item responses, the items exhibited a different set of characteristics when grouped together as sources of self-efficacy (Table 10). In all sources, mean scores were above the midpoint of their possible ranges. For example, the range of possible total scores for mastery experience was 8-800. The midpoint would have been 396, but the reported mean was 614.88. This condition was an indication of negative skewness resulting from most participants rating their self-perceptions highly. There were differences, however, in the range of the participants' responses as indicated by the variances associated with each source of self-efficacy. Mastery experience had the greatest amount of variability (*SD* = 122.88) while physiological state had the least (*SD* = 87.98). This finding may be related to the number of items reflecting each source of self-efficacy. Since there was a greater number of items

reflecting mastery experience (*n* =8) than vicarious experience (n = 5), there was a greater chance that participants would have different self-perceptions of those experiences.

Table 10
Descriptive Statistics for Sources of Self-Efficacy (N = 290)

Sources of Self-Efficacy	Number of items	*M*	*SD*	Min.	Max.
Mastery Experience	8	614.88	122.88	193	800
Vicarious Experience	5	361.40	91.41	6	500
Verbal/Social Persuasion	6	480.64	101.72	132	600
Physiological State	5	394.97	87.98	24	500
MPSES composite	24	1851.89	328.60	821	2400

Correlations were used as another descriptive statistic to examine the responses. The inter-item correlations on the MPSES ranged from a slight negative correlation (*r* = -.04) between item 15 ("I am learning, or have learned, to control nervousness during a performance.") and item 18 ("I have watched other students of similar musical ability as me... and then decided whether I could ... perform the same piece of music.") to a moderately strong correlation (*r* = .66) between item 17 ("Performing on my instrument makes me feel good.") and item 23 ("I enjoy participating in musical performances."). The lack of correlation between items 15 and 18 was not surprising given that each item reflected a different source of self-efficacy. Item 15 reflected physiological state and item 18 reflected vicarious experience. The strong correlation between items 17 and 23 could

have been anticipated since both items reflected the same source of self-efficacy. Most inter-item correlations fell within the .10 to .40 range suggesting modest to moderate relationships between the item responses (Appendix N). This finding confirmed that the items were not measuring the same self-perception, but measured various facets of self-efficacy.

Correlations were also used to evaluate the functioning of the items within each source of self-efficacy. The relationships between the items and the sources of self-efficacy are reported in Table 9. This examination indicated moderate relationships between most items and the source of self-efficacy. These moderate relationships were acceptable since the items were constructed to measure a range of self-perceptions related to self-efficacy. The items associated with verbal/social persuasion demonstrated stronger relationships with the source than items associated with vicarious experience. This finding was consistent with prior research that identified the challenges of constructing effective items to measure vicarious experience (Lent, Lopez, & Bieschke, 1991; Lopez & Lent, 1992; Matsui, Matsui, & Ohnishi, 1990; Usher & Pajares, 2006). The relationship between item 19 and physiological state (r = .26) was also a concern. The low correlation of item 19 with physiological state indicated it may not be measuring the same construct as the other items. As mentioned above, item 19 also had a low mean score.

After assessing the inter-item correlations and the correlations between the items and the sources of self-efficacy, I evaluated the relationships between the items and the composite construct of self-efficacy (Table 9). These item-total correlations provided additional information about the functioning of the items. In this evaluation, item 19 (“I

do not worry about making mistakes...") demonstrated the weakest relationship with the total score (r = .25) and item 23 ("I enjoy participating in musical performances.") exhibited the strongest relationship (r = .62). Although responses to item 19 demonstrated a low mean score, weak relationship with its source of self-efficacy physiological state, and weak relationship with the composite construct, I found no benefit from removing it from the analysis. Item 19 remained in the analysis.

The correlations among the sources of self-efficacy and with the composite construct are reported in Table 11. Among the sources of self-efficacy, mastery experience and verbal/social persuasion had the strongest relationship (r = .68) while vicarious experience and physiological state exhibited the weakest relationship (r = .34). These moderate correlations suggested that the sources of self-efficacy on the MPSES were measuring different, yet related, dimensions of self-efficacy. If all of the relationships had been very strong (e.g., r = .90), the sources may have represented a single factor. This finding supported Bandura's belief that the sources of self-efficacy contributed unique forms of information to the overall construct of self-efficacy. In comparison to self-efficacy in other subject areas, it is interesting to note that the rank order of the strengths of the correlations between the sources of self-efficacy and the composite construct were the same as the rank order found in other studies (1st – mastery experience, 2nd – verbal/social persuasion, 3rd – physiological state, and 4th – vicarious experience).

Table 11
Correlations between the Sources of Self-Efficacy and Composite Self-Efficacy (N =290)

Sources	Mastery	Vicarious	Verbal/Social	Physiological
Mastery	-			
Vicarious	.54	-		
Verbal/Social	.68	.44	-	
Physiological	.60	.34	.60	-
Composite Self-Efficacy	.89	.71	.85	.77

All correlation were significant at the $p < .001$ level.

Analytical Assumptions

Analytical procedures are based on mathematical models. Therefore, the data must meet certain requirements. This study used procedures such as confirmatory factor analysis, multivariate analysis of variance, and regression to examine the data. Since these procedures are based on the general linear model, the data used in these procedures must meet similar requirements, or assumptions, for the analytical procedures to function properly. The primary assumptions for these procedures are independence of observations, normal distribution, and homogeneity of variance. The MANOVA requires the data to meet an additional assumption of homogeneity among the covariance matrices. If the data do not meet these assumptions, consequences must be considered when interpreting the results.

The first assumption was independence of the observations. Although the participants were nested within classes, the observations were considered independent. The questionnaires were self-report measures and completed by participants on separate computers. The topics addressed by the questionnaires were not related to specific treatments received within the classes. Common experiences were shared by the

participants, but those experiences may be interpreted in different ways based on the personal history and characteristics of the participants. In addition, bivariate scatter plots of responses to the sources of self-efficacy were not elliptical indicating randomness in all relationships except between mastery experience and verbal/social persuasion.

The next assumption was that the observations were normally distributed. According to Glass and Hopkins (1996), meeting this assumption renders the means and variances statistically independent. In other words, there is no correlation between mean and standard deviation values. Since multivariate analyses were used in this investigation, a discussion of univariate normality will be included within the discussion of multivariate normality. Byrne (2005) suggested several methods to establish multivariate normality: (a) univariate distributions must be normal, (b) joint distributions must be normal, and (c) all bivariate scatter plots must be linear and homoscedastic (i.e., each variable having equal variance).

Skewness and kurtosis values indicated all univariate distributions were normal with the exception of physiological state (Table 12). Physiological state exhibited excessive negative skewness ($\gamma_1 = -1.30$). The joint distribution of the music self-efficacy composite score, however, suggested a normal distribution of scores and showed no adverse effects from the negative skewness of physiological state. In an alternate test of normality, the bivariate scatterplot of mastery experience and verbal/social persuasion created an elliptical pattern. In contrast, other scatterplots appeared less well-defined. Scatter plots with random patterns provide evidence of independence but not normality. Along with this evidence of univariate non-normality, multivariate skewness (*b1p*) and kurtosis (*b2p*) were also non-normal, $b1p = 5.33$, χ^2 (20, $N = 290$) = 261.52, $p < .05$ and

b2p = 32.79, 95% CI [10.28, 10.80]. Stevens (2009), however, stated that the *F* statistic was robust to violations of normality in ANOVA and MANOVA procedures, but the consequence of non-normality was a reduction in power.

Table 12
Data Distribution Indicators for All Scales (N =290)

Dependent Variable	Skewness	Kurtosis
Mastery Experience	-0.81	0.50
Vicarious Experience	-0.75	0.61
Verbal/Social Persuasion	-0.95	0.33
Physiological State	-1.30	1.95
Music Self-Efficacy (composite score)	-0.70	0.12
Mathematics Self-Efficacy[a]	-0.58	-0.38
Music Self-Esteem[b]	-0.32	-0.16
Music Aptitude[c]	0.04	-0.41

[a] n = 289
[b] n = 286
[c] n = 222

The homogeneity of variance among group responses was another underlying assumption. A comparison of the standard deviations between middle school (MS) and high school (HS) participants revealed that middle school participants consistently exhibited greater variance in responses than high school participants (Mastery experience: MS *SD* = 140.12, HS *SD* = 101.08; vicarious experience: MS *SD* = 98.01, HS *SD* = 81.88; verbal/social persuasion: MS *SD* = 107.32, HS *SD* = 95.66: and physiological state: MS *SD* = 98.47, HS *SD* = 75.37). Since score distributions were negatively skewed, the Brown-Forsythe test was used for determining the homogeneity of variance between groups. The Brown-Forsythe test is appropriate for evaluating non-normal distributions because it bases its comparisons on medians rather than means. One

limitation of the Brown-Forsythe test, however, is that it only makes one-way comparisons. In other words, it tests one independent variable at a time. In a comparison by grade level, results indicated that the variances were homogenous for verbal/social persuasion and physiological state ($p > .05$) but not for mastery experience and vicarious experience ($p < .05$). Stevens (2009) stated, "As long as the group sizes are approximately equal (largest/smallest < 1.5), *F* is robust" (p. 227). In this comparison, the groups were almost equal in size with 150 middle school participants and 140 high school participants. Consequently, the unequal variances between the groups on mastery experience and vicarious experience had little influence on the results.

In a comparison by ensemble, results from the Brown-Forsythe procedure indicated that variances were not homogenous for mastery experience, but they were homogenous for the other sources of self-efficacy. Group sizes in this comparison, however, were unequal. There were 87 band participants, 153 choral participants, and 50 string orchestra participants. Stevens (2009) recommended, "When the large variances are associated with the large group sizes, then the *F* statistic is *conservative*" (p. 227, emphasis in original). The word conservative meant that when the proposed level of significance was $\alpha = .05$, the actual level of significance was less than .05. In this study, choral participants (the largest group) exhibited the greatest amount of variance for each source of self-efficacy. Stevens concluded that this violation would also result in a reduction in power.

In MANOVA, the homogeneity of variance assumption extends to homogeneity of the covariance matrices. Box's M test was used to evaluate the homogeneity of the covariance matrices between grade levels and among ensembles. Results indicated the

matrices were not homogenous between grade levels, χ^2 (10, N = 290) = 20.45, p < .05, and among ensembles, χ^2 (20, N = 290) = 39.35, p < .05. Since the requirements of this assumption are difficult to meet, Stevens (2009) provided assistance reporting that Monte Carlo simulations found the MANOVA was robust to violations of the homogeneity of variance and homogeneity of covariance assumptions as long as sample sizes were similar. Additional requirements included the data being of interval scale and a large sample size ($N \geq 200$). In this study, the group sizes were unequal between ensembles, but the data were interval level and a large sample size was used (N = 290). Given the unequal sample sizes and violations of assumptions, Pillai's Trace was used as the outcome statistic from the MANOVA procedure rather than Wilks' lambda. Pillai's Trace is more conservative statistic and consequently more robust to violations of assumptions.

Research Questions

Question 1. *Do responses of secondary school music student to items on the Music Performance Self-Efficacy Scale adequately fit Bandura's theoretical model of self-efficacy?*

Since the MPSES was based on Bandura's theoretical model, confirmatory factor analysis (CFA) was an appropriate procedure to evaluate the relationships between the items (i.e., variables) and the sources of self-efficacy (i.e., latent factors). CFA compares an a priori factor structure to the factor structure of the collected data. It is a five step procedure in which the model is specified, identified, estimated, evaluated, and modified. The identified model is presented in Figure 3. The model was identified by setting the loading to 1.0 from each source of self-efficacy to one of its reflective items. The loading was set to 1.0 between mastery experience and item 1, between vicarious experiences and

item 2, between verbal social persuasion and item 3, and between physiological state and item 15. This portion of the model is referred to as the 1st order. The loadings between the composite construct of self-efficacy and the sources of self-efficacy is referred to as the 2nd order. In the 2nd order, the loading between self-efficacy and mastery experience was also set to 1.0. Maximum likelihood estimation was then used to estimate the other loadings, errors, and variances.

CFA results were contradictory. On one hand, the χ^2 test and the comparative fit index (CFI) identified a lack of fit between the data and Bandura's proposed model, χ^2 (248, N = 290) = 568.09, $p < .05$, and CFI = 0.84. On the other hand, the standardized root-mean-square residual (SRMR) and the root-mean-square error of approximation (RMSEA) indicated close to adequate fit, SRMR = .06 and RMSEA = .07. These interpretations were based on Hu and Bentler's (1998) benchmarks for identifying good fit, CFI $\geq$ .95, SRMR $\leq$ .08, and RMSEA $\leq$ .06.

Since the χ^2 statistic is sensitive to sample size, it was not surprising to find misfit given the size of this sample. Larger sample sizes increase the χ^2 statistic resulting in a rejection of the null hypothesis. The discrepancy between the CFI and the SRMR and RMSEA may have been a reflection of the difference in the type of indices they represent. The CFI is an incremental index in which the model formed by the collected data is compared to a restricted baseline model. The SRMR and RMSEA are absolute indices that assess "how well the a priori model reproduces the sample data" (Hu & Bentler, 1998, p. 426). The lack of fit indicated by the low CFI value may have resulted from correlations between the items, errors, or factors. The restricted baseline model does not allow for correlations. Consequently, correlated errors in the data model resulted in

misfit with the baseline model. Being absolute indices, the SRMR and RMSEA had a greater capacity to accommodate correlations.

An examination of the modification indices identified correlations between items and residuals. I was interested in improving the fit between the data model and the proposed model, but did not want to deviate from Bandura's theoretical framework or the principles upon which the items where developed. Therefore, I avoided incorporating correlations between items, and loading items onto different or multiple sources of self-efficacy. I was, however, willing to include correlations between errors. Even though errors are the unpredictable part of the score variance, it is conceivable that there may be relationships between errors, because the items were closely related. In the modification indices, I found correlations between the errors associated with items 19 and 15, items 12 and 9, and items 4 and 1. After incorporating these correlations into the model, there was some improvement in the fit indices, χ^2 (245, N = 290) = 501.62, $p < .05$, CFI = .87, SRMR = .06, RMSEA = .06. The CFI exhibited the greatest improvement in fit, from .84 to .87. This improvement provided evidence that the misfit indicated by the CFI may have been more of a reflection of an incremental index's sensitivity to correlations than misfit between the models.

In the same article, Hu and Bentler (1998) recommended using the SRMR as the primary fit index and supplementing it with findings from another index. The RMSEA was one of their secondary suggestions. Based on the evidence of good fit provided from the SRMR and RMSEA values, I propose that the MPSES generated data that adequately fit Bandura's model. In the 1st order, the items accurately reflected the four sources of

self-efficacy, and in the 2nd order, the sources of self-efficacy accurately reflected the composite construct of self-efficacy.

Question 2. *How much influence do mastery experience, vicarious experience, verbal/social persuasion, and physiological state have on self-efficacy in music performance among secondary school music students?*

The confirmatory factor analysis produced estimated loadings that indicated the strength of the individual items (i.e., variables) to reflect the latent constructs (i.e., factors). I reported standardized loadings to allow for comparisons. In this analysis, the 1st order loadings ranged from .20 for item 19 to .83 for item 23 (Table 13). One interpretation of this difference is that item 23 was stronger in its ability to reflect physiological state than was item 19.

Table 13
Estimated Standardized Loadings of Sources of Self-Efficacy onto Items

Mastery Experience	StdYX	Vicarious Experience	StdYX	Verb./Soc. Persuasion	StdYX	Phys. State	StdYX
Item 1	.60	Item 2	.55	Item 3	.69	Item 15	.34
Item 4	.40	Item 5	.54	Item 7	.59	Item 17	.78
Item 6	.61	Item 11	.42	Item 9	.57	Item 19	.20
Item 8	.37	Item 18	.43	Item 13	.57	Item 23	.83
Item 10	.61	Item 20	.39	Item 21	.64	Item 24	.68
Item 12	.56			Item 22	.61		
Item 14	.44						
Item 16	.58						

Similarly, I used CFA to determine the estimated loadings of self-efficacy on the sources of self-efficacy. The loadings in rank order were mastery experience (StdYX =

.98), verbal/social persuasion (StdYX = .91), physiological state (StdYX = .83), and vicarious experience (StdYX = .75). This finding suggested that mastery experience had greater similarities with the composite construct of self-efficacy than the other sources and made the greatest contribution. The rank order of these loadings was similar to findings in other studies (Lent, Lopez, & Bieschke, 1991; Lopez & Lent, 1992; Matsui, Matsui, & Ohnishi, 1990; Usher & Pajares, 2006, 2009).

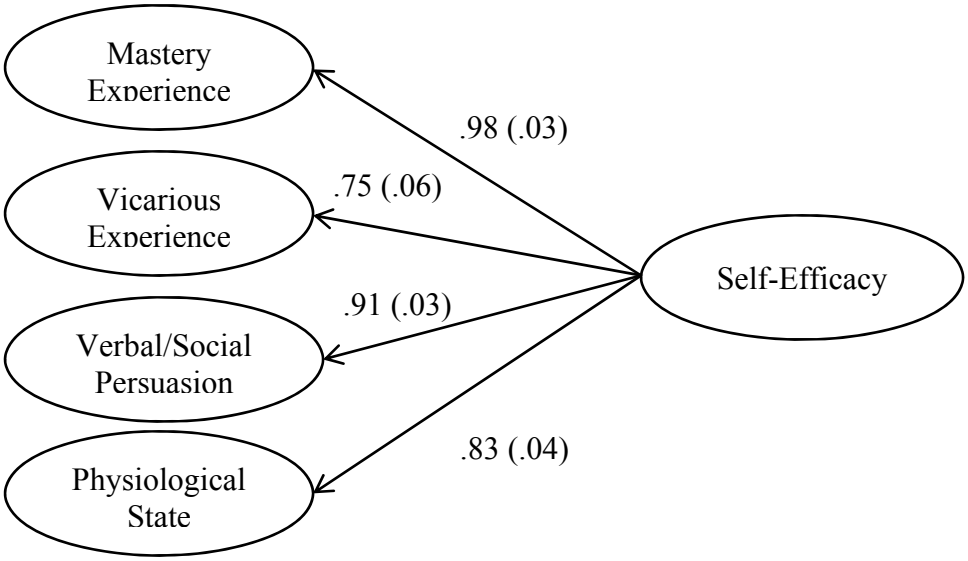

Figure 4. Standardized Loadings and Standard Errors from Self-Efficacy onto the Sources of Self-Efficacy

Question 3. *Are the items on the Music Performance Self-Efficacy Scale invariant for middle and high school students?*

CFA was also used to answer this question. To determine factor invariance, various parameters were systematically constrained to determine whether there were differences in the responses across groups of participants while ability was held constant. Initially, separate CFAs were run for the independent variable grade level (i.e., middle

and high school). The resulting χ^2 statistics and degrees of freedom were then added together to form a configural, or baseline model (Table 14). As mentioned above, the distribution of data for the MPSES was not multivariate normal. Therefore, the Satorra-Bentler scaled chi-square difference was used to determine if the change from one model to the next was significant or not.

Next, the loadings between the sources of self-efficacy and the items were constrained to be the equal for each grade and a CFA was run. Initial results did not indicate a significant difference between the $SBS\Delta\chi^2$ and the Δdf , but the CFI dropped from .79 to .77. Dimitrov (2010) suggested that a change in CFI > -1.0 indicated misfit in the model. I examined the modification indices and identified a possible misfit between item 1 and mastery experience. By allowing item 1 ("I have had positive experiences performing music in the past.") to vary freely while the other loadings were constrained, the fit of the model improved. The ΔCFI was only from .79 to .78 and the $SBS\Delta\chi^2$ was from 37.78 to 27.20. The implication of this decision was that only partial measurement invariance could be achieved. This finding also implied that middle and high school participants responded differently to item 1. It is also important to note that allowing item 1 to vary freely reduced the Δdf from 20 to 19.

After the metric invariance testing process, the data were evaluated for scalar invariance. Scalar invariance testing involved a process similar to metric invariance testing, but in this process, the intercepts were constrained to be equal. The results of this test indicated no significant difference between the models. There was not a large enough change in the $SBS\Delta\chi^2$ value from the previous configuration when compared to the Δdf to suggest that the items were biased in favor of either group. Generally speaking, there

must be more than a 2:1 ration between χ^2 and the degrees of freedom to indicate a significant difference.

The next step in factor invariance testing was to evaluate the invariance of the item uniquenesses. This process involved constraining the error variances and covariances associated with each item to be equal. In this comparison, there was a significant difference between the $SBS\Delta\chi^2$ (63.65) and the Δdf (24). Another indicator of misfit was the ΔCFI (-.04). This finding suggested that the error variances and covariances across the items were not the same for middle and high school participants.

In addressing the research question, there was evidence of strong partial measurement invariance. All items, except for item 1, demonstrated metric and scalar invariance. Middle and high school participants responded similarly to the items resulting in similar loadings for each group and similar intercepts (i.e., mean scores). These findings supported the appropriateness of using the MPSES with middle and high school students.

Table 14

Testing for Factor Invariance among 24 MPSES Items across Middle (n = 150) and High School (n = 140)Music Students

Model	χ^2	$SB\chi^2$[a]	*df*	$SBS\Delta\chi^2$[b]	Δ*df*	CFI	ΔCFI	SRMR	RMSEA
Baseline	946.46	715.49	492			.79[c]		0.08	0.08
Metric Invariance[d]	982.37	744.03	511[e]	27.20	20	.78	-.01	0.08	0.09
Scalar Invariance[f]	1021.66	781.11	531[g]	37.42	20	.77	-.01	0.08	0.09
Invariance of Item Uniquenesses[h]	1134.32	853.17	555	63.65	24	.73	-.04	0.12	0.09

[a] The Satorra-Bentler chi-square was used in calculating the SBSΔχ².

[b] The Satorra-Bentler scaled chi-square difference (SBSΔχ²) was used to determine the change in chi-square in order to accommodate the non-normal multivariate distribution of MPSES scores.

[c] The CFI value reported here represents the fit of 1st order model, while the CFI value reported in other parts of the study represents the fit of the 2nd order model.

[d] Metric invariance — equal factor loadings across groups.

[e] Theoretically, the degrees of freedom should have changed by 24 since 24 loadings were constrained. Four loadings, however, were previously constrained to 1.0 to identify the model, and one loading (mastery experience by item 1) was allowed to remain "free" since it was a primary source of misfit when constrained as identified in the modification indices.

[f] Scalar invariance — equal intercepts across groups.

[g] Theoretically, the degrees of freedom should have changed by 24 since 24 intercepts were constrained. The error variances of the four latent factors, however, were fixed to 0.0 in the previous models. Once the intercepts were constrained, it was no longer necessary to fix the error variances and these four parameters were added back into the degrees of freedom.

[h] Invariance of item uniqueness — equal item error variances/covariance across groups.

Questions 4 and 5. *Are there differences in the value attributed to mastery experience, vicarious experience, verbal/social persuasion, and physiological state by band, chorus, and string orchestra students? Are there differences in the value attributed to mastery experience, vicarious experience, verbal/social persuasion, and physiological state by middle and high school music students?*

Multivariate analysis of variance (MANOVA) was used to answer questions four and five. MANOVA is an extension of the univariate analysis of variance (ANOVA) and used to compare group means on multiple dependent variables. One fundamental difference between the two procedures is that ANOVA uses scalars, or integers, in its linear model while MANOVA uses matrices. When comparing multiple variables, MANOVA has two primary advantages over using multiple ANOVA procedures. Those advantages are being more effective in controlling Type I error and having the ability to incorporate correlations among the variables into the analytical model. In this study, MANOVA was used to compare differences among grade levels, types of ensembles, and the interaction between grade and ensemble based on scores from four dependent variables (i.e., the four sources of self-efficacy).

The dependent variables in this analysis were the sources of self-efficacy as measured by the subsections of the MPSES—mastery experience, vicarious experience, verbal/social persuasion, and physiological state. The item responses from each source were totaled and recorded as a unique score for each participant. For example, the responses to items 1, 4, 6, 8, 10, 12, 14, and 16 were added together to form a mastery experience score for each participant. These aggregated scores were then used to represent each source of self-efficacy in the analysis.

As mentioned above, the data did not meet the required assumptions for using MANOVA. The data were not normally distributed, the variances were not homogenous, the covariance matrices were not homogenous, and the cell sizes were unequal (Appendix O). There were several rationales, however, that supported the use of MANOVA: (a) the observations were independent; (b) the sample size was fairly large; (c) the data were continuous and at the interval level; and (d) the primary consequence was a reduction in power. With those violations in mind, I approached the analyses conservatively using Pillai's Trace in determining whether to reject, or fail to reject, the null hypotheses.

The MANOVA was conducted as a 3x2 factorial design using SAS software. The PROC GLM procedure was initiated with ensemble and grade level entered in the class statement (i.e., independent variable). Mastery experience, vicarious experience, verbal/social persuasion, and physiological state were entered in the model statement (i.e., dependent variable). The results of the MANOVA indicated no significant differences among ensembles, grade levels, and the interaction of ensemble and grade level (Table 15). This finding provided evidence to suggest that music students in band, chorus, and string orchestra ensembles at the middle and high school levels attribute similar values to the composite construct of self-efficacy.

Table 15
MANOVA Results from Music Performance Self-Efficacy Scale (N = 290)

Variable	V^a	*F*	*Num df*	*Den df*	*p*
Ensemble	0.0388	1.39	8	564	NS
Grade	0.0244	1.76	4	281	NS
Ensemble*Grade	0.0522	1.89	8	564	NS

[a] Pillai's Trace

Although not mandatory when reporting non-significant results, I included the univariate results to facilitate discussion. The univariate tests compared differences among groups on single dependent variables. Since this study was exploratory in many ways, reporting the univariate results may be useful in future research. Means and standard deviations are reported in Appendix O. The findings of the univariate *F*-tests were not as uniform as the multivariate test. For mastery experience, the interaction between ensemble and grade level was significant (Table 16). The mean score of participants from one grade level and one type of ensemble were significantly different than the mean score of participants from the other grade level and another type of ensemble. For vicarious experience, there were significant differences in all comparisons of the ANOVA (Table 17). There were significant differences in the two-way model: the main effects of ensemble and grade level, and the interaction between ensemble and grade level. These differences were an indication that participants in various types of ensembles and grade levels place diverse values on vicarious experiences.

Table 16
Univariate F-test for Mastery Experience (N = 290)

Source	*df*	*SS*	*MS*	*F*	*p*
Model	5	133858	26772	1.80	NS
Error	284	4230156	14895		
Corrected Total	289	4364013			
Ensemble	2	83922[a]	41961	2.82	NS
Grade	1	32665[a]	32665	2.19	NS
Ensemble*Grade	2	90493[a]	45246	3.04	<.05

[a]Type III SS

Table 17
Univariate F-test for Vicarious Experience (N = 290)

Source	*df*	*SS*	*MS*	*F*	*p*
Model	5	148618	29724	3.72	<.05
Error	284	2266421	7980		
Corrected Total	289	2415040			
Ensemble	2	56290[a]	28145	3.53	<.05
Grade	1	51364[a]	51364	6.44	<.05
Ensemble*Grade	2	86429[a]	43215	5.42	<.05

[a]Type III SS

There were no significant differences for verbal/social persuasion (Table 18) or physiological state (Table 19). Pairwise comparisons were also conducted using the Tukey method with an adjusted Bonferroni alpha level of .01. No significant differences were found in the pairwise comparisons.

Table 18
Univariate F-test for Verbal/Social Persuasion (N = 290)

Source	*df*	*SS*	*MS*	*F*	*p*
Model	5	71876	14375	1.40	NS
Error	284	2918139	10275		
Corrected Total	289	2990015			
Ensemble	2	47294[a]	23647	2.3	NS
Grade	1	14345[a]	14345	1.40	NS
Ensemble*Grade	2	55150[a]	27575	2.68	NS

[a]Type III SS

Table 19
Univariate F-test for Physiological State (N = 290)

Source	*df*	*SS*	*MS*	*F*	*p*
Model	5	46387	9277	1.20	NS
Error	284	2190531	7713		
Corrected Total	289	2236918			
Ensemble	2	40314[a]	20157	2.61	NS
Grade	1	18316[a]	18316	2.37	NS
Ensemble*Grade	2	9810[a]	4905	0.64	NS

[a]Type III SS

Question 6. *To what extent does music aptitude predict self-efficacy in music performance?*

A regression equation was constructed to predict self-efficacy in music performance using music aptitude scores. Scores from Gordon's AMMA (1989) (Appendix F) were used as the predictor variable music aptitude and scores from the MPSES (Appendix B) were used as the criterion variable self-efficacy. Only scores from those participants who completed the AMMA and the MPSES (n = 222) were included in this analysis. To improve comparisons with other regression coefficients, the scores were standardized (i.e., transformed into z scores) prior to constructing the equation. The mean music aptitude score (47.79) was subtracted from each participant's music aptitude score and then divided by the standard deviation (10.07) to obtain a z score for music aptitude. In the same way, the mean self-efficacy score (1827.77) was subtracted from each participant's music self-efficacy score and divided by the standard deviation (363.39) to obtain a z score for self-efficacy in music performance. Transforming the raw scores into

z scores will allow future comparisons to be made between the results in this study with standardized scores in other studies.

Results indicated that music aptitude scores predicted a modest increase in self-efficacy scores, $\beta = .16(.07)$ (Figure 5). This relationship was found to be significant, $F(1, 220) = 6.12$, $p < .05$, 95% CI [.03, .29]. In addition, approximately 3% of the variance in self-efficacy was found to be predictable by music aptitude. This finding inferred that students with higher levels of music aptitude tended to have somewhat higher levels of self-efficacy. Scores from the AMMA in this study were consistent with findings in previous studies. Participants in this study included a combination of middle and high school students. Their scores on the AMMA ($M = 47.79$, $SD = 10.07$) were almost exactly between scores produced by a group of high school students ($M = 50.60$, $SD = 7.91$) (Gordon, 1989) and scores produced by a group of middle school students ($M = 46.10$, $SD = 8.01$) (Gordon, 2004) in previous studies. These comparisons provided evidence that the AMMA scores from this study were consistent with AMMA scores in other studies.

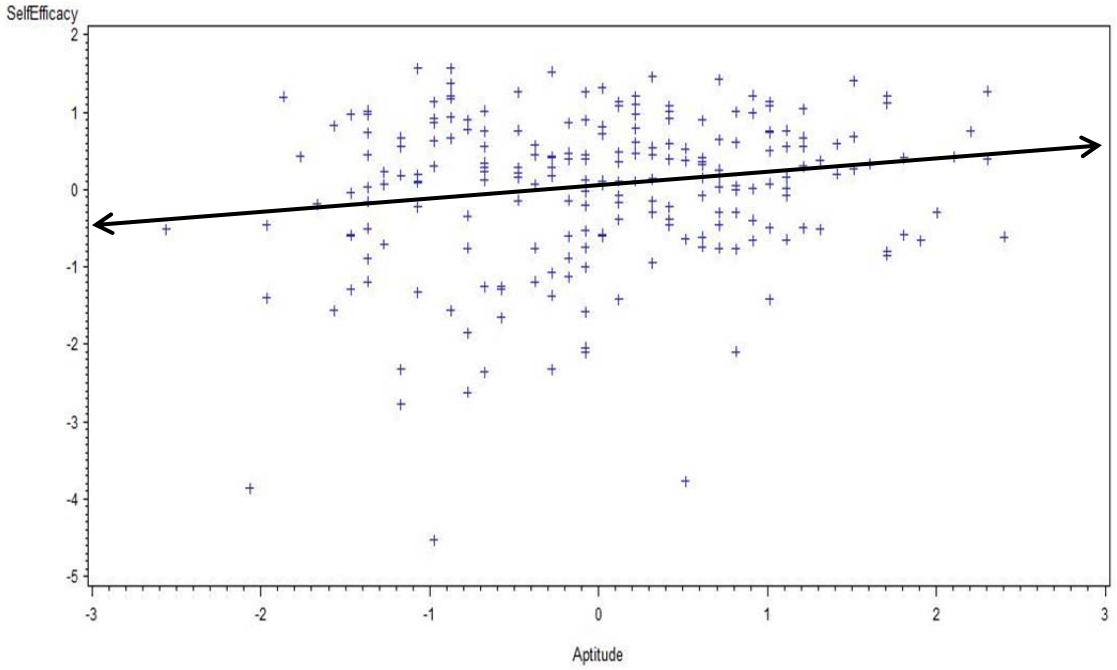

Figure 5. Regression Results Predicting Music Self-Efficacy from Music Aptitude. Estimated regression line (β = .16) with music aptitude and music self-efficacy scores in standard deviation units.

Question 7. *To what extent does evidence support the use of the Music Performance Self-Efficacy Scale as a measure of self-efficacy in music performance?*

This study followed recommendations proposed by the American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education (1999) to determine the validity and reliability of scores generated by a measurement instrument. In this case, the study sought to examine evidence of whether the MPSES produced results that were representative of self-efficacy in music performance. This process involved establishing evidence of test content,

response processes, internal structure, relations to other variables, test-criterion relationships, and consequences of testing. Due to the design of this study, it was not possible to establish evidence based on test-criterion relationships and consequences of testing. Establishing evidence of test-criterion relationships would have required the study to include a measure of music performance achievement. Measures of music performance can be highly subjective and differ greatly between teachers, schools, and school districts. Since this study intended to include a diverse population of participants, incorporating preexisting music performance measures such as grades may have provided data that was invalid or unreliable. Since the time required to complete the questionnaires ranged from 30-60 minutes, I believed that adding an additional instrument to measure music performance would have further reduced response rates. Therefore, I decided not to include evidence that would have been used to establish the test-criterion relationship. In another form of evidence, consequences of testing were not addressed because this was a "one-shot" study. There would be no contact with participants after completing the initial questionnaires. Establishing evidence of test-criterion relationships and consequences of testing are important issues that were beyond the scope of this study.

To establish evidence of test content, a panel of music education experts reviewed the scale items to determine whether the items elicited responses that accurately reflected self-perceptions of music performance in secondary school settings. The panel consisted of four professors of music education and seven doctoral students of music education. The panel made several recommendations that were incorporated into the scale. Those recommendations included:

- Specify number of performers in items that reference small and large ensembles.

- Create new item to address solo performance.
- Create new item to include experiences of watching someone other than a peer or professional perform music.
- Remove wording in item 2 "who are similar to me in some way perform well."
- Remove item "I have improved my music performance skills by watching other students who are similar to me in some way perform well."

The panel also made several recommendations that I did not incorporate into the scale. Those recommendations were:

- Include a response of "not applicable."
- Change the response format from 1-100 to 1-10.

Since this study was an effort to fill the gap in self-efficacy research between other subject areas and music education, I based this scale on models in other subject areas. No other scales included a response option of "not applicable" and the 1-100 response format had been shown to have psychometric advantages (Byrne, 2005; Pajares, Hartley, & Valiante, 2001; Usher & Pajares, 2009). After the revisions, the panel agreed the scale items reflected the four sources of self-efficacy and accommodated Bandura's request for items that varied in generality, strength, and level.

According to Bandura, self-efficacy is a subjective self-perception and not an objective measurement such as grades or achievement test scores. The MPES was designed to be a self-report measure. To evaluate the accuracy of the response process, self-efficacy scores from the MPSES were correlated with teacher ratings of student self-efficacy. Teachers (n = 8) provided ratings on a scale of 1 (very low) to 5 (very high) of their students' (n = 175) self-efficacy beliefs based upon their observations of the

students' music behavior in the classroom. The result of this correlation ($r = .42$, $p < .001$) indicated a moderate relationship between the self-efficacy scores and the teachers' ratings. This finding provided evidence that the self-report responses generated by the MPSES were similar to the students' observable behavior. It is essential to acknowledge in interpreting this correlation that there were extreme differences in the correlations between student responses and ratings among the teachers. These correlations ranged from $r = -0.09$ for one teacher to $r = 0.66$ for another. Given that information, the response process used by the MPSES was reasonably accurate in extracting self-perceptions of self-efficacy from the participants.

Findings from the first research question provided evidence that the internal structure of the MPSES was consistent with Bandura's theoretical framework. These findings indicated that there was adequate fit between the collected data and the proposed model, SRMR = .06 and RMSEA = .06. The MPSES generated scores that were consistent with the underlying theoretical framework. In addition, the rank order of the loading strengths was consistent with Bandura's proposed model and findings in other studies. Mastery experience demonstrated the strongest loading, followed by verbal/social persuasion, physiological state, and vicarious experience.

Evidence of relationships to other variables was examined from discriminant and convergent perspectives. From a discriminant perspective, a weak relationship was found between scores on the MPSES and scores on the Sources of Middle School Mathematics Self-Efficacy Scale (SMSMSES) ($r = .29$, $p < .001$). This relationship suggested that there was some common element between the two, but the MPSES was measuring a different construct than the SMSMSES. Although both measures assessed self-efficacy,

the MPSES was able to discriminate between self-efficacy in music performance and self-efficacy in mathematics. From a convergent perspective, scores on the MPSES exhibited a moderately strong relationship with scores on the Self-Esteem of Musical Ability ($r = .67$, $p < .001$). The strength of this relationship provided evidence that results from the MPSES were consistent with results from SEMA. Self-efficacy and self-esteem are similar constructs and consequently similar scores were obtained on the two measures. In addition, the correlation between the teachers' ratings and the students' self-efficacy scores ($r = .42$) provided evidence of multi-method convergence. The teachers' ratings were based on observable behaviors while the student scores were generated from a self-report questionnaire.

Summary

Middle and high school participants ($N = 290$) in band, chorus and string orchestra classes completed four questionnaires via the Internet. The participants were from 10 schools in the southeast and western regions of the United States. Results from confirmatory factor analysis indicated that responses collected from the Music Performance Self-Efficacy Scale fit Bandura's proposed model, χ^2 (245, $N = 290$) = 501.62, $p < .05$, CFI = .87, SRMR = .06, RMSEA = .06. This conclusion was based on the SRMR and RMSEA values. From this analysis, the standardized loading between self-efficacy and mastery experience was found to be the strongest (StdYX = .98), followed by verbal/social persuasion (StdYX = .91), physiological state (StdYX = .83), and vicarious experience (StdYX = .75). Factor invariance testing suggested that the items on the MPSES were functioning without bias among middle and high school participants when the loadings and intercepts were constrained to be equal. However, the

equality among the groups was not tenable when the error variances were constrained. Item 1 was an exception to this finding eliciting different responses from each group.

In a 3x2 MANOVA, no difference was found by ensemble, grade level, or the interaction of ensemble and grade level. Data from items reflecting the four sources (i.e., mastery experience, vicarious experience, verbal/social persuasion, and physiological state) were aggregated into separate sources and treated as separate dependent variables.

An examination of the relationship between music aptitude and self-efficacy in music performance identified a standardized regression coefficient of .16 for music aptitude suggesting that self-efficacy in music performance increased by .16 standard deviation units (58.14 raw score points) for each standard deviation unit increase in music aptitude. In addition, music aptitude accounted for 3% of the variance in self-efficacy.

Finally, evidence was established to support the use of the MPSES as a valid and reliable measure of self-efficacy in music performance. Evidence of test content was established by a panel of music education experts. Evidence of response process was established by the correlation of MPSES scores and teacher ratings of student self-efficacy (r = .42). Evidence of internal structure was established by confirmatory factor analysis (SRMR = .06, RMSEA = .06). Responses from the MPSES fit Bandura's proposed model and the loadings of the composite construct self-efficacy onto the sources of self-efficacy exhibited the same rank order as described in Bandura's theory and found in other studies. Mastery was strongest, followed by verbal/social persuasion, physiological state, and vicarious experience. Evidence of relationships to other variables was established from a discriminant perspective by the modest correlation (r = .29) between scores on the MPSES and the Sources of Middle School Mathematics Self-

Efficacy Scale, and from a convergent perspective by the strong relationship between the MPSES and the Self-Esteem of Musical Ability (r = .67). In addition, the correlation between teacher ratings of student self-efficacy and participant scores on the MPSES (r = .42) provided evidence of multi-method convergence. Results from a test-retest procedure (r = .87) and the internal consistency of responses from the MPSES (α = 88) provided evidence of reliability.

Chapter 5: Summary, Discussion, and Recommendations

Summary

Although often overlooked in music instruction, self-efficacy in music performance has been shown to play an important role in music education. Previous studies have identified self-efficacy as the best predictor of music performance achievement when compared to intrinsic value, general self-efficacy, and self-regulation (McCormick & McPherson, 2003; McPherson & McCormick, 2000). It was also identified as a mediating variable between other variables such as formal practice and the outcome variable of performance achievement (McPherson & McCormick, 2006). The current study extended these findings by investigating the sources of self-efficacy and their contribution to the composite construct of self-efficacy. The findings of this study along with previous work in this area have created a theoretical path that begins with the sources of self-efficacy and ends with music performance achievement.

Results of this study indicated that data collected by a measure of self-efficacy in music performance (i.e., Music Performance Self-Efficacy Scale) fit Bandura's proposed model. This finding provides a theoretical connection between self-efficacy in music performance and self-efficacy in other subject areas—they share a common framework. The implications of this finding are numerous. Along with good fit, no differences were found by ensemble or grade level among secondary music students. The sources of self-efficacy functioned consistently across ensembles and grade level. In a unique inquiry, a relationship was also identified between music aptitude and self-efficacy in music

performance. This finding brings the role of self-efficacy into question when determining the relationship between music aptitude and music achievement.

Finally, evidence was established to support the validity and reliability of scores from the MPSES. An accurate measure of self-efficacy in music performance enables researchers to continue investigations of self-efficacy in music performance and provides educators with a diagnostic instrument to monitor their students' self-efficacy development.

Discussion

The purpose of this study was to develop a greater understanding of self-efficacy in music performance. It sought to (a) contribute to fundamental knowledge of self-efficacy in music performance, (b) determine whether or not the Music Performance Self-Efficacy Scale was valid and reliable, and (c) provide insights for developing self-efficacy among secondary school band, chorus, and string orchestra students. The discussion section below addresses the first two objectives while the third objective is addressed in the recommendations for music education section.

Fundamental knowledge.

Bandura's theoretical framework as described in his social cognitive theory was found to be applicable to the self-efficacy beliefs held by middle and high school students in band, chorus, and string orchestra classes. Results obtained from a diverse group of participants ($N = 290$) demonstrated good fit with Bandura's proposed model. Good fit was expressed by the SRMR and RMSEA values (.06 and .06, respectively). A lack of fit, however, was expressed by the CFI value (.87) and continued to be a concern. Based on Hu and Bentler's (1989) recommendation, I used the SRMR index as my primary

indicator of fit and the RMSEA as a secondary indicator. These indices were found to be more robust to violations of assumptions than other fit indices.

The finding of good fit confirmed the use of Bandura's framework for studying self-efficacy in music performance and provides a link between self-efficacy in music performance and self-efficacy in other subject areas. Based on this finding, music education researchers can create models for music education based on preexisting models in other subject areas. Although the content of music education is unique, the self-perceptions of instructional experiences in music function in ways similar to the ways they function in other subject areas. Models such as Lent and others' (1994) career intention model can be replicated in the subject area of music and provide the foundation for further understanding of these psychological constructs and their impact on students.

Most participants in the study exhibited high levels of self-efficacy as represented by negative skewness among the sources of self-efficacy and the composite construct (Table 12). This finding, however, may not have been an accurate representation of the total population of music students. Since participants were required to submit signed parent permission forms, the results may be more a reflection of a population that has the motivation and ability to return those forms than of all music students. Therefore, the results of this study must be approached with caution and not generalized to a larger population. This issue is being brought to light in order to frame the interpretation of the results. Based on the responses from these students, the relative contribution of each source of self-efficacy was consistent with findings from other subject areas. Mastery experience exerted the greatest influence, followed by verbal/social persuasion, physiological state, and finally vicarious experience.

Educators should be aware of the mastery experiences they incorporate into their instruction since these experiences exert a strong degree of influence on their students' self-perceptions of musical ability. In this study, data were collected from students who were presently enrolled in music classes. Item 8 stated "I have had positive experiences performing simple music." Within this group of students, item 8 had the highest mean score (M = 89.19) and least amount of variance (SD = 16.98). These results indicated that the participants had a solid background of mastery experiences. This finding runs contrary to practices in some music classes in which educators prioritize the development of psychomotor skills and technical proficiency. Although beyond the scope of this study, this discussion begs the question: how would students who were no longer enrolled in music classes respond to item 8? I propose that developing mastery experiences is one strategy for curbing the high dropout rate among secondary school music students. Students benefit from performing simpler pieces by building confidence in their abilities and establishing memories of successful music performances.

Verbal/social persuasion also exerted strong influence on self-efficacy. This statement was supported by the strength of the standardized loading between self-efficacy and verbal/social persuasion (StdYX = .91) and the correlation between the two constructs (r = .85). Positive reinforcement of accomplishment through verbal and social praise builds self-efficacy. Although this finding may appear as common sense, educators should consider using a systematic approach to ensure that their students are verbally and socially supported for their efforts. This support can come from the teacher and others whose opinions are valued by the student. In addition, the persuasion must be authentic. Wise and Trunnell (2001) found that positive reinforcement has greater impact on self-

efficacy when preceded by mastery experiences. The reinforcement must be interpreted by the student as coming from something they have done, rather than coming from something they could do.

The other sources of self-efficacy, physiological state and vicarious experience, should not be ignored. Standardized loadings and correlations were moderately high between these sources and self-efficacy (Physiological state: StdYX = .83, r = .77; Vicarious experience: StdYX = .75, r = .71). Educators should consider the physiological state of their students while the students are performing music. Does performing music make the students feel anxious and uncomfortable? Based on the results of item 19, "I do not worry about small mistakes during a performance," (M = 58.32, SD = 34.59), I have concluded that most students worry about making mistakes and this worry contributes to uncomfortable physiological states. Therefore, instruction addressing performance anxiety may benefit the students and educators may wish to reevaluate the learning environment of their classrooms. Strategies that render music performance as an enjoyable and exciting experience should be implemented.

Finally, vicarious learning experiences can become part of an instructional routine. Students can be taught to make constructive comparisons of themselves to others. Instruction that helps them identify specific characteristics and develop a greater understanding of the preparation performed prior to their observation can help overcome their initial response of "I could never do that." As noted in the literature, there has been speculation on the differences between peer and adult vicarious experiences. I believe these differences can be assimilated into this source of self-efficacy as subcategories that

reflect the generality of this source. Vicarious experiences come from a variety of individuals such as peers, professionals, family members, and others.

Another important finding was that no significant difference was found in self-efficacy among students in different grade levels and different ensembles. When compared by grade level, music students in middle school (i.e., grades 6, 7, and 8) and high school (i.e., grades 9, 10, 11, and 12) reported similar levels of music performance self-efficacy. The implication of this finding is that self-efficacy functions in the same way among students of different ages (i.e., grade level) and different music performance activities (i.e., ensembles). Consequently, educators can expect to find strategies created to strengthen self-efficacy in one type of ensemble or grade level, to deliver similar results in others. This finding also highlighted another difference between self-efficacy and self-esteem. Randles (2011) reported that self-esteem for music decreased with age. This study, however, found that self-efficacy remained the same. This difference may have resulted from music and non-music students being included in Randles' study, while only music students were included in this study. This issue may need to be reexamined.

The ability of music aptitude to predict self-efficacy in music performance was examined in this study. Scores from the Advanced Measures of Music Audiation (AMMA) predicted a .16 increase in the standardized scores from the Music Performance Self-Efficacy Scale (i.e., 58.14 raw-score points). Music aptitude also accounted for 3% of the variance in music performance self-efficacy. These findings may be of interest to researchers investigating music aptitude or self-efficacy in music performance. More importantly however, these findings link two bodies of literature—studies associated with Gordon's (1989) concept of music aptitude, and studies associated with Bandura's (1986)

concept of self-efficacy. I am not aware of another connection made between these two bodies of literature prior to this study.

Validity and reliability of the Music Performance Self-Efficacy Scale.

The findings of this study supported the use of the MPSES as a valid and reliable instrument to assess self-efficacy among secondary school music students. Several forms of evidence were considered in making this judgment. Prior to administering the study, evidence of content validity was established through the evaluation of experts in the field of music education. These experts found the scale items to accurately reflect the four sources of self-efficacy as represented in music instruction among secondary school music students. They also believed the items represented sufficient generality, strength, and level as described by Bandura (2006). A comparison of scores generated by the MPSES were moderately correlated (r = .42) with teacher ratings of student self-efficacy demonstrating the effectiveness of the response process. Evidence based on the internal structure of the scale was established through confirmatory factor analysis. The CFA produced results that indicated the MPSES generated data that adequately fit Bandura's theoretical model and the relative strengths of the loadings were consistent with theory and findings from other studies. Factor invariance testing was used to establish that items on the MPSES were unbiased and did not favor students in a particular type of ensemble.

Finally, the MPSES was found to produce scores that demonstrated evidence of discriminant and convergent validity. The relationship between MPSES scores and scores from Usher and Pajares' (2009) Sources of Middle School Mathematics Self-Efficacy Scale were modest (r = .29). Conversely, the MPSES was found to produce scores that converged with similar constructs. The relationship between MPSES scores and scores

from Schmitt's (1979) Self-Esteem of Musical Ability were moderately strong (r = .67). These findings provided evidence of the relationships between scores on the MPSES and scores reflecting other variables. In addition, evidence of reliability was established through a test-retest procedure (r =.87) and a high level of internal consistency (α = .88). Based on these forms of evidence as recommended by the American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education (1999), scores from the MPSES can be interpreted as being valid and reliable.

The results from this study were different in several ways from results in the preliminary study (Zelenak, in press). This finding does not come as a surprise since many characteristics of the studies were different. First, participants from the first study were middle school students from one school while participants in the current study were middle and high school students from 10 schools. Second, the preliminary study was delivered to the students by the researcher in a paper-and-pencil format while the current study was administered online. Third, the items were grouped together by source of self-efficacy in the preliminary study while they were arranged in random order in the current study. In my opinion, I believe the third difference may have exerted the most influence on the difference in scores. The high degree of internal consistency in the first study may have been a result of item order that was not present in the current study.

Recommendations

Initial research of self-efficacy was conducted by psychologists interested in determining the effectiveness of their treatments. Soon afterward, educators became aware of the relevance of Bandura's framework in educational settings. Most self-

efficacy studies were conducted in academic subjects such as mathematics, science, and language arts. There has been a growing awareness, however, of self-efficacy in non-academic subjects such as physical education (Feltz, Short, & Sullivan, 2008). The recommendation in the first section are aimed at encouraging music education researchers to continue this investigation. The following section was included for music educators. There are practical applications for findings in this study and these recommendations are an attempt to bridge the gap between theory and practice.

For further research.

As pointed out by McCormick and McPherson (2003), it is surprising how little research has been conducted on this topic of self-efficacy in music performance given the findings of the relationships between self-efficacy in music performance and music achievement. The current study was conceived as an extension of previous work in this area. Previous work has provided insights in the relationships of music performance self-efficacy and other variables, but there has been little investigation of the sources that contribute to the development of these self-efficacy beliefs. This study has provided important insight into the sources of self-efficacy in music performance.

This study also provided a unique contribution to the field by examining the relationship between music aptitude and self-efficacy in music performance. The modest relationship between the two constructs provides a rationale for future studies that may connect two previously disparate areas of study—music aptitude as presented in Gordon's music learning theory and self-efficacy as presented in Bandura's social cognitive theory. Connecting the work of Fullen and Gordon to the work of McPherson and McCormick may provide multidimensional understandings of the underlying

processes in music learning. Music aptitude, self-efficacy, and music achievement may be constructs in one dimension that also have relationships with variables in another dimension such as grade level, self-regulation, formal practice, and informal practice. Understanding the relationships between these constructs is essential to improving the quality of music education.

Future research may build upon these foundations in several ways. First, findings from the follow-up univariate tests suggested that there may be differences in mastery experience based on the interaction of ensemble and grade level. This finding raises the question of whether there is a difference in the contribution of mastery experience between students in middle school chorus and high school band, or some other combination of students. Also, there is a need to investigate the influence of vicarious experience in the music classroom. Given trends in education toward more cooperative learning, this investigation would provide information in determining whether vicarious experiences with peers, adults, or other types of musicians are more effective in developing self-efficacy beliefs.

A second recommendation for future research would be to expand this investigation to other populations. Given the findings in this study, the MPSES would be a suitable instrument to measure the self-efficacy beliefs of undergraduate musicians and musicians in community ensembles. The instrument would not be suitable for use among elementary aged students although investigations of self-efficacy in music among younger students is an equally important topic.

Third, there is a need to conduct research in self-efficacy that utilizes a performance outcome measure. McCormick and McPherson (2003) and McPherson and

McCormick (2000, 2006) were successful in identifying the relationships between self-efficacy and music performance using graded music performance evaluations such as the Trinity College examination and the Australian Music Examinations Board performance evaluation. The current study needs to be replicated and extended to include a performance measure. An evaluation of the relationship between the sources of self-efficacy and performance achievement would provide valuable information.

A fourth and final recommendation for further research would be to conduct an experimental study with a self-efficacy component as the treatment. In this study, one group of students would be provided with increased opportunities to develop their sources of self-efficacy while a control group would receive their regular music instruction. Since most of the previous studies of self-efficacy have been based on descriptive designs, an experimental design might provide other useful information from a different perspective. In addition, the study might include a qualitative component to capture students' individual thoughts and concerns.

For music educators.

Self-efficacy in music performance is one aspect of music instruction that has been greatly overlooked in the past. In this study, I sought to develop fundamental knowledge about self-efficacy and also raise an awareness of this psychological construct among music educators. I believe one reason for the lack of attentiveness to this construct has been the absence of an instrument to measure it among music students. This belief led me to develop the Music Performance Self-Efficacy Scale. The total score generated by the MPSES provides a general measure of self-efficacy in music performance. The difference in this scale from other measures of music self-efficacy is that it was designed

to measure the *sources* of self-efficacy that contribute to the overall construct. Although the total score provides some useful information, the subtotal scores may be more useful for teachers because they can be used to diagnose the strengths and weaknesses in the sources of self-efficacy among their students. The results of this diagnosis can then be used to make instructional decisions. Teachers can determine if their students are in need of reinforcement in (a) mastery experiences, (b) opportunities to observe others perform music, (c) positive feedback, or (d) strategies to control performance anxiety. When using the MPSES, the mean scores presented in this study are offered as general benchmarks with which teachers can compare their students' scores (Table 20) and not national normalized scores. Since no significant differences were found among grade levels or ensembles, teachers can use the same set of mean scores for middle or high school students.

Table 20
Results from the Music Performance Self-Efficacy Scale by Ensemble (N = 290)

Sources of Self-Efficacy	Number of items	*M*	Per Item Mean	*SD*	Min.	Max.
Mastery Experience	8	614.88	76.86	122.88	8	800
Vicarious Experience	5	361.40	72.28	91.41	5	500
Verbal/Social Persuasion	6	480.64	80.11	101.72	6	600
Physiological State	5	394.97	78.99	87.98	5	500
MPSES total	24	1851.89	77.16	328.60	24	2400

The following is a hypothetical example of the way in which a teacher may use the scores from the MPSES. The items for each source of self-efficacy are items 1, 4, 6, 8, 10, 12, 14, and 16 for mastery experience; items 2, 5, 11, 18, and 20 for vicarious experience; items 3, 7, 9, 13, 21, and 22 for verbal/social persuasion; and items 15, 17, 19, 23, and 24 for physiological state. A middle school choral teacher administers the MPSES to his or her choral students. Mean scores can be calculated using common spreadsheet software such as Excel (Appendix M), or statistical software such as SPSS or SAS. She finds that the students had an overall mean score of 1620, a mean score of 590 in the mastery experience section, a mean score of 190 in the vicarious experience section, a mean score of 420 in the verbal/social persuasion section, and a mean score of 300 in the physiological state section. By comparing their scores to the findings of this study, a teacher can approximate the relative strength or weakness of mastery experience, vicarious experience, verbal/social persuasion, and physiological state of his or her student. In this example, the students scored low in the vicarious experience section. The teacher would then be alerted that the students needed more opportunities to watch peers and adults perform music and gain greater understanding of the possible accomplishments within their grasp.

Once student scores have been established and evaluated, changes in instruction may be required. A few suggestions follow. Mastery experience is built upon previous successes. Bandura (1997) reminded us, "Successes build a robust belief in one's personal efficacy. Failures undermine it, especially if failures occur before a sense of efficacy is firmly established" (p. 80). This statement is an important reminder to educators of beginning music students. Students are easily discouraged when first starting

out and it is essential to provide opportunities for success. Students also need to be presented with strategies for overcoming musical challenges. Rehearsal strategies provide the music student with a sense of control over these challenges and the mastery of difficult tasks raises beliefs in one's abilities. In addition, mastery experiences are developed over time. The longer students persist at an activity while experiencing success, the stronger and more stable their beliefs will become.

Vicarious experiences also contribute to self-efficacy beliefs. Individuals make judgments about themselves based on comparisons with others. Receiving ratings higher than group norms increases self-efficacy beliefs while ratings at or below an average lowers those beliefs regardless of the level of achievement. They also compare themselves to others engaged in similar activities and who have characteristics similar to themselves. Bandura (1997) stated, "The greater the assumed similarity, the more pervasive are the models' successes and failures" (p. 87). Individuals who are uncertain about their abilities are more sensitive to vicarious experiences. Therefore, the vicarious experiences of beginning students need to be monitored. Even though students may be of the same age and grade level, their level of experience may promote unfair comparisons.

Verbal/social persuasions contribute almost as much influence on self-efficacy beliefs as mastery experience. In the current study, the first order factors of mastery experience (.98) and verbal/social persuasion (.91) demonstrated strong loadings on the second order factor of self-efficacy. In the music classroom, positive feedback and social experiences are essential to forming self-efficacy beliefs. Bandura (1997) proposed that positive feedback initiates a series of occurrences that ultimately increase achievement. He wrote, "persuasive boosts in perceived efficacy lead people to try hard enough to

succeed, self-affirming beliefs promote development of skills and a sense of personal efficacy" (p. 101). He also cautioned that the positive feedback must be perceived by the individual as authentic and false praise imparts a negative influence on the individual and discredits the persuader. In this spirit, Wise and Trunnell (2001) found that verbal feedback was more effective when following a mastery experience than when following a vicarious experience.

Physical and emotional states also influence self-efficacy beliefs. In music, performance anxiety can be debilitating. Stress can result from anticipation of poor performance or impending lack of control. In some cases, physical symptoms are indicators of uncertain self-beliefs. The development of coping strategies and increased mastery experiences are some ways to address performance anxiety.

It is important to use student scores from the MPSES appropriately. These scores are not intended to be measures of achievement and therefore should never be used in calculating grades. They should only be used by teachers to identify strengths and weaknesses among their students and drive instruction. Scores from the MPSES can also be used by teachers as feedback on the effectiveness of their instruction.

List of References

Abel, J. L, & Larkin, K. T. (1990). Anticipation of performance among musicians: Physiological arousal, confidence, and state-anxiety. *Psychology of Music, 18,* 171-182.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Austin, J. R., & Vispoel, W. P. (1998). How American adolescents interpret success and failure in classroom music: Relationships among attributional beliefs, self-concept, and achievement. *Psychology of Music, 26*, 26-45.

Aydin, Y. C., & Uzuntiryaki, E. (2009). Development and psychometric evaluation of the high school chemistry self-efficacy scale. *Educational and Psychological Measurement, 69*(5), 868-880.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, *37*(2), 122-147.

Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Bandura, A. (1990). *Multidimensional scales of perceived academic efficacy*.

Stanford, CA: Stanford University.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman and Co.

Bandura, A (2006). *Self-efficacy beliefs of adolescents.* Information Age Publishing.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588-606.

Berry, J. M., & West, R. L. (1993). Cognitive self-efficacy in relation to personal mastery and goal setting across the lifespan. *International Journal of Behavioral Development, 16,* 351-379.

Betz, N. E., & Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational Behavior, 23*, 329-345.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen *&* J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Byrne, B. M. (1989). Multigroup comparisons and the assumption of equivalent construct validity across groups: Methodological and substantive issue. *Multivariate Behavioral Research, 24,* 503-523.

Byrne, B. M. (2005). Factor analytic models: Viewing the structure of an assessment instrument from three perspectives. *Journal of Personality Assessment, 85,* 1, 17-32.

Choi, N., Fuqua, D. R., & Griffin, B. W. (2001). Exploratory analysis of the structure of scores from the Multidimensional Scales of Perceived Self-Efficacy. *Educational and Psychological Measurement, 61*(3), 475-481.

Craske, M. G., & Craig, K. D. (1984). Musical performance anxiety: The three-systems model and self-efficacy theory. *Behavioral Research and Therapy, 22*(3), 267-280.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory.* New York: Harcourt Brace Jovanovich College Publishers.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281-302.

Dalgety, J., Coll, R. K., & Jones, A. (2003). Development of chemistry attitudes and experiences questionnaire. *Journal of Research in Science Teaching, 40*(7), 649-668.

Dimitrov, D. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development, 43,* 121-149.

Feltz, D. L., Short, S. E., & Sullivan, P. J. (2008). *Self-Efficacy in sport.* Champaign, IL: Human Kinetics.

Fouad, N. A., & Smith, P. L. (1996). A test of a social cognitive model for middle school students: Math and science. *Journal of Counseling Psychology, 43*(3), 338-346.

Fullen, D. L. (1993). *An investigation of the validity of the Advanced Measures of Music Audiation with junior high and senior high school students* (Doctoral dissertation) Retrieved October 17, 2010, from ProQuest Dissertations & Theses: Full Text (AAT 9316479).

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction.* New York: Pearson Education, Inc.

Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology.* Needham Heights, MA: Allyn & Bacon.

Gordon, E. E. (1989) *Advanced Measures of Music Audiation.* Chicago, IL: GIA Publications, Inc.

Gordon, E. E. (1993). *Learning sequences in music: Skill, content, patterns: A music learning theory.* Chicago, IL: GIA Publications, Inc.

Gordon, E. E. (2004). *Continuing studies in music aptitudes.* Chicago, IL: GIA Publications, Inc.

Greenberg, M. (1970). Musical achievement and the self-concept. *Journal of Research in Music Education, 18*(1), 57-64.

Hayward, C. M., & Gromko, J. E. (2009). Relationships among music sight reading and technical proficiency, spatial visualization, and aural discrimination. *Journal of Research in Music Education, 57*(1), 26-36.

Hendricks, K. S. (2009). Relationships between the sources of self-efficacy and changes in competence perceptions of music students during an all-state orchestra event (Doctoral dissertation. Retrieved October 17, 2010, from ProQuest Dissertations & Theses: Full Text (AAT 3362920).

Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and Evaluation in Counseling and Development, 34,* 177-189.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3,* 424-453.

Joo, Y-J., Bong, M., & Choi, H-J. (2000). Self-efficacy for self-regulated learning, academic self-efficacy, and Internet self-efficacy in web-based instruction. *Educational Technology, Research, and Development, 48*(2), 5-17.

Kendall, M. J. (1988). Two instructional approaches to the development of aural and instrumental performance skills. *Journal of Research in Music Education, 36*(4), 205-219.

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior, 45,* 79-122.

Lent, R. W., Lopez, F. G., & Bieschke, K. J. (1991). Mathematics self-efficacy: Sources and relation to science-based career choice. *Journal of Counseling Psychology, 38*(4), 424-430.

Lent, R. W., Lopez, F. G., Brown, S. D., & Gore, P. A. (1996). Latent structure of the sources of mathematics self-efficacy. *Journal of Vocational Behavior, 49,* 292-308.

Lopez, F. G., & Lent, R. W. (1992). Sources of mathematics self-efficacy in high school students. *Career Development Quarterly, 41*(1), 3-9.

Matsui, T., Matsui, K., & Ohnishi, R. (1990). Mechanisms underlying math self-efficacy learning of college students. *Journal of Vocational Behavior, 37,* 225-238.

McCormick, J., & McPherson, G. (2003). The role of self-efficacy in a musical performance examination. *Psychology of Music, 31*(1), 37-51.

McPherson, G. E., & Hendricks, K. S. (2010). Students' motivation to study music: The United States of America. *Research Studies in Music Education, 32,* 201-213.

McPherson, G. E., & McCormick, J. (2000). The contribution of motivational factors to instrumental performance in a music examination. *Psychology of Music, 31*(1), 37-51.

McPherson, G. E., & McCormick, J. (2006). Self-efficacy and music performance. *Psychology of Music, 34*(3), 322-336.

Messick, S. (1989). Validity. In R. L. Linn (Ed), *Educational measurement* (3rd ed., pp 13-103). New York: Macmillan.

Nielsen, S. W. (2004). Strategies and self-efficacy beliefs in instrumental and vocal individual practice: A study of students in higher music education. *Psychology of Music, 32*(4), 418-431.

Nunnally, J. C. (1967). *Psychometric theory*. New York: McGraw-Hill.

Pajares, F. (1996). Role of self-efficacy beliefs in the mathematical problem solving of gifted students. *Contemporary Educational Psychology, 21,* 325-344.

Pajares, F. (2007). Empirical properties of a scale to assess writing self-efficacy in school contexts. *Measurement and Evaluation in Counseling and Development, 39,* 239-249.

Pajares, F., Hartley, J., & Valiante, G. (2001). Response format in writing self-efficacy

assessment: Greater discrimination increases prediction. *Measurement and Evaluation in Counseling and Development, 33*, 214-221.

Pajares, F., Miller, M. D., & Johnson, M. J. (1999). Gender differences in writing self-beliefs of elementary school students. *Journal of Educational Psychology, 91*(1), 51-60.

Pajares, F., & Valiante, G. (1999) Grade level and gender differences in the writing self-beliefs of middle school students. *Contemporary Educational Psychology, 24,* 390-405.

Papageorgi, I., Hallam, S., & Welch, G. F. (2007). A conceptual framework for understanding musical performance anxiety. *Research Studies in Music Education, 28,* 83-107.

Pastorelli, C., Caprara, G. V., Barbaranelli, C., Rola, J., Rosza, S., & Bandura, A. (2001). The structure of children's perceived self-efficacy: A cross-national study. *European Journal of Psychological Assessment, 17*(2), 87-97.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire.* National Center for Research to Improve Post Secondary Teaching and Learning Project on Instructional Processes and Educational Outcomes, document HE 024 935.

Randles, C. (2010). The relationship of compositional experiences of high school instrumentalists to music self-concept. *Bulletin of the Council for Research in Music Education, 184,* 7-18.

Randles, C. (2011). “What is a good musician?” An analysis of student beliefs. *Arts Education Policy Review, 112,* 1-8.

Ritchie, L., & Williamon, A. (2007). Measuring self-efficacy in music. Presented at the International Symposium on Performance Science, 2007. Published by AEC. ISBN 978-90-9022484-8.

Rohwer, D., & Polk, J. (2006). Practice behaviors of eighth-grade instrumental musicians. *Journal of Research in Music Education, 54*(4), 350-362.

Schmitt, M. C. J. (1979). Development and validation of a measure of self-esteem of musical ability (Doctoral dissertation). Retrieved October 17, 2010, from ProQuest Dissertations & Theses: Full Text (AAT 8009164).

Schunk, D. H. (1983). Developing children's self-efficacy and skills. The roles of social comparative information and goal setting. *Contemporary Educational Psychology, 8,* 76-86.

Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist, 25,* 71-86.

Sehman, K. H. (2000). The effects of breath management instruction on the performance of elementary brass players. *Journal of Research in Music Education, 48*(2), 136-150.

Sherer, M., Maddux, J. E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. W. (1982). The self-efficacy scale: Construction and validation. *Psychological Reports, 51,* 663-671.

Skinner, B. F. (1953). *Science and human behavior.* New York: Macmillan.

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences.* New York: Routledge, Taylor, & Francis Group.

Stewart, J. L. (2002). *Factors related to students' decisions to continue in band.* (Doctoral dissertation). University of Texas at Austin. UMI #3110589.

Svengalis, J. N. (1978). *Music attitude and the preadolescent male* (Doctoral dissertation). Retrieved October 17, 2010, from ProQuest Dissertations & Theses: Full Text (AAT 7902953).

Thompson, B. (2004). *Exploratory and confirmatory factor analysis.* Washington, D.C.: American Psychological Association.

Usher, E. L., & Pajares, F. (2006). Sources of academic and self-regulatory efficacy beliefs of entering middle school students. *Contemporary Educational Psychology, 31,* 125-141.

Usher, E. L., & Pajares, F. (2009). Sources of self-efficacy in mathematics: A validation study. *Contemporary Educational Psychology, 34,* 89-101.

Wehr-Flowers, E. L. (2007). *An exploratory model of jazz self-efficacy and gender* (Doctoral dissertation). Retrieved October 17, 2010, from ProQuest Dissertations & Theses: Full Text (AAT 3281418).

Williams, D. A. (2007). What are music educators doing and how well are we doing it? *Music Educators Journal, 94*(1), 18-23.

Wise, J. B., & Trunnell, E. P. (2001). The influences of sources of self-efficacy upon efficacy strength. *Journal of Sport and Exercise Psychology, 23,* 268-280.

Zelenak, M. S. (2009). An exploratory study of middle school music students' self-efficacy and intention to enroll in future music classes in traditional and technology-enhanced learning environments. *Research Perspectives in Music Education, 13,* 38-46.

Zelenak, M. S. (in press). Development and validation of the Sources of Music Performance Self-Efficacy Scale. *Music Education Research International.*

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*(1), 51-59.

Appendices

Appendix A

Demographic Information

Welcome! You have reached the Self-Efficacy in Music Performance questionnaire site.

Privacy Protection: Only the principal investigator will have access to the information you provide in this survey. You will not be asked for your name, but will be asked for other information that will be used to compare groups of students.

Please tell me about yourself.

What is your birth date? (mm/dd/yyyy)

_____

What is your gender?

_____Male _____Female

What is your race?

____White____Black____Asian____Hispanic___Native American
____Indian____Mixed____Other

What is your grade level?
____6. ____7. ____8. ____9. ____10. ____11. ____12.

In what music class did you turn in your parent permission for and gain access to this study?
____Band ____Chorus ____String Orchestra

Who is your music teacher for this class?

_____

Appendix B

Music Performance Self-Efficacy Scale

Directions: Respond to the following statements based on your current level of musical ability, experience, and primary instrument/voice. There are no right or wrong answers. Indicate to what degree you either agree or disagree with the statement by inputting **any whole number** between **1 (Strongly Disagree)** and **100 (Strongly Agree)** into the box. Carefully consider the number you choose.

1, 2, 3, 4.....10.....20.....30.....40.....50.....60.....70.....80.....90.....97, 98, 99, 100
Strongly Disagree — Strongly Agree

Hint: Do not press the "Enter" key after entering your answers. Use your mouse to move between questions, or the "Tab" key. If you do hit Enter, you will receive warning signs. Continue to answer the questions. The warnings will disappear when you move to the next section.

_____1. I have had positive experiences performing music in the past (Choose a number between 1-100).

_____2. I have improved my music performance skills by watching professional musicians perform well.

_____3. My friends think I am a good performer on my primary instrument/voice.

_____4. I have had positive experiences performing in large ensembles (more than 11 performers)

_____5. I have improved my music performance skills by watching someone I know perform well (parent, brother, sister, church member, etc,).

_____6. I have had positive experiences performing music solo.

_____7. Members of my family believe I perform well.

_____8. I have had positive experiences performing simple music.

_____9. People have told me that my practice efforts have improved my performance skills.

_____10. I have had positive experiences performing complicated music.

_____11. I have used other music students as models to improve my performance skills.

_____12. I have overcome musical challenges through hard work and practice.

_____13. I have received positive feedback on music performance evaluations.

_____14. I have used a practice routine to help me prepare for my performances.

_____15. I am learning, or have learned, to control my nervousness during a performance.

_____16. I have had positive experiences performing music in a small ensemble (2-10 performers).

_____17. Performing with instrument/voice makes me feel good.

_____18. I have watched other students with similar music ability as me perform a piece of music, and then decided whether I could, or could not, perform the same piece of music.

_____19. I do not worry about making small mistakes during a performance.

_____20. I have compared my performance skills with those of other students who are similar in musical ability to me.

_____21. My music teacher has complimented me on my musical performance.

_____22. I have met or exceeded other people's expectations of being a good musician for someone my age.

_____23. I enjoy participating in musical performances.

_____24. I have positive memories of most, or all, of my past music performances.

Appendix C
(From Preliminary Study)

Identification Code:_____

Music Performance Self-Efficacy Scale

Directions: Respond to the following statements based on your current level of musical ability, experience, and primary instrument or voice. There are no right or wrong answers. Indicate to what degree you either agree or disagree with the statement by writing ***any*** number between
0 **(Strongly Disagree)** and ***100*** **(Strongly Agree)** on the line next to the statement. Carefully consider the number you choose.

0	**10**	**20**	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**100**
Strongly Disagree										Strongly Agree

Part I - (Mastery experiences)

_____1. I have had positive experiences performing music in the past.

_____2. I have had positive experiences performing in large ensembles.

_____3. I have had positive experiences performing solo, or, in a small ensemble.

_____4. I have had positive experiences performing simple music.

_____5. I have had positive experiences performing complicated music.

_____6. I have overcome musical challenges through hard work and practice.

_____7. I have used a practice routine to help me prepare for my performances.

Part II - (Vicarious experiences)

_____8. I have improved my music performance skills by watching professional musicians, who are similar to me in some way, perform well.
_____9. I have improved my music performance skills by watching other students, who are similar to me in some way, perform well.
_____10. I have used other music students as models to improve my performance skills.

_____11. I have compared my performance skills with those of other students who are similar in musical ability to me.
_____12. I have watched other students of similar musical ability as me perform a piece of music, and then decided whether I could, or could not, perform the same piece of music.

Part III - (Verbal/Social persuasion)

_____13. My friends think I am a good performer on my primary instrument.

_____14. Members of my family believe I perform well.

_____15. My music teacher has complimented me on my musical performance.

_____16. People have told me that my practice efforts have improved my performance skills.
_____17. I have received positive feedback on music performance evaluations.

_____18. I have met or exceeded other people's expectations of being a good musician for someone of my age.

_____19. Write only the number 9 for this answer (not 0-100 rating).

Part IV - (Physiological state)

_____20. Performing with my instrument makes me feel good (Return to using 0-100 rating).

_____21. I enjoy participating in musical performances.

_____22. I am learning, or have learned, to control nervousness during a performance.

_____23. I do not worry about small mistakes during a performance.

_____24. I have positive memories of most, or all, of my past musical performances.

Place an "X" next to the correct response.

What is your gender? _____Male _____Female

What is your race? _____Asian ____ Black _____Hispanic ____ Indian(American)
_____White ____ Mixed _____Other

What is your grade level? _____6 ____ 7 ____ 8
____ 9 ____ 10 ____11 ____12

Are you enrolled in a music class at this school? Yes_____ No_____

If you answered "Yes," what is the name of the class?_____

If you answered "Yes," what is your primary instrument?_____

What is the teacher's name in whose class you took this survey?_____

Appendix D

Sources of Middle School Mathematics Self-Efficacy Scale*

*Adapted from "Sources of self-efficacy in mathematics: A validation study" by E. L. Usher and F. Pajares (2009), *Contemporary Educational Psychology, 34*, 89-101.

There are no right or wrong answers. Indicate to what degree you either agree or disagree with the following statements by **inputting an whole number between 1(Strongly Disagree) and 100 (Strongly Agree) in the box.** Carefully consider the number you choose.

1, 2, 3, 4.....10.....20.....30.....40.....50.....60.....70.....80.....90.....97, 98, 99,100
Strongly Disagree — Strongly Agree

_____25. I make excellent grades on math tests.

_____26. I have always been successful with math.

_____27. Even when I study very hard, I do poorly in math.

_____28. I got good grades in math on my last report card.

_____29. I do well on math assignments.

_____30. I do well on even the most difficult math assignments.

_____31. Seeing adults do well in math pushes me to do better.

_____32. When I see how my math teacher solves a problem, I can picture myself solving the problem in the same way.

_____33. Seeing kids do better than me in math pushes me to do better.

_____34. When I see how another student solves a math problem, I can see myself solving the problem in the same way.

_____35. I imagine myself working through challenging math problems successfully.

_____36. I compete with myself in math.

_____37. My math teachers have told me that I am good at learning math.

____38. People have told me that I have a talent for math.

_____39. Adults in my family have told me what a good math student I am.

_____40. I have been praised for my ability in math.

_____41. Other students have told me that I'm good at learning math.

_____42. My classmates like to work with me in math because they think I'm good at it.

_____43. Just being in math class makes me feel stressed and nervous.
_____44. Doing math work takes all of my energy.

_____45. I start to feel stressed-out as soon as I begin my math work.

_____46. My mind goes blank and I am unable to think clearly when doing math work.

_____47. I get depressed when I think about learning math.

_____48. My whole body becomes tense when I have to do math.

Appendix E

Self-Esteem of Musical Ability*

*Adapted from *Development and validation of a measure of self-esteem of musical ability,* (Doctoral dissertation) by M. C. J. Schmitt, 1979. Retrieved from ProQuest dissertations and theses, full text (AAT 8009164). Copyright 1979 by M. C. J. Schmitt.

These are statements about your musical ability. Some may be true for you; others may not. Read each statement then select strongly disagree, disagree, agree, or strongly agree.

_____49. I can read music well.

_____50. I know music well enough to help others learn it.

_____51. I could write music if I got a little help.

_____52. Leading others in singing or playing would be difficult for me.

_____53. I can play or sing difficult rhythms.

_____54. I think I could win a music contest if I really tried.

_____55. I am glad when asked to sing or play for others.

_____56. Music teachers often embarrass me by asking questions which I can't answer.

_____57. I am glad that my parents expect a lot from me in music.

_____58. I usually feel uncomfortable when I am around music.

_____59. I feel good when my parents notice my progress in music and praise me.

_____60. When I stop and think about it, I really believe I'm talented in music.

_____61. My parents believe I can learn to play or sing really well.

_____62. Teachers notice my progress in music and give me credit.

_____63. My friends really like my playing or singing.

_____64. My music teachers don't expect much of me.

_____66. When music teachers leave the room, I can take their place.

_____67. I feel that kids would laugh at me if I made mistakes in music.

_____68. I'd be glad if teachers asked me to play or sing for programs.

_____69. I believe I could become a professional singer or player.

_____70. Kids notice my progress in music and give me credit for what I can do.

_____71. Music is all right for others but not for me.

_____72. I will not be good enough to be in choir or band in college.

_____73. I enjoy playing for others.

_____74. I expect a lot of myself in music.

_____75. People my age admire my musical ability.

_____76. I have such a rough time in music class that I often feel worthless.

_____77. I find myself helping my friends with their music.

_____78. I am glad my family likes to listen to me perform music.

_____79. Music is harder for me than for the other kids.

_____80. I am not satisfied with my progress in music.

_____81. At least one member of my family says I am really good in music.

_____82. I would like to have a professional career in music.

_____83. Usually I enjoy practicing music.

_____84. I like it when music teachers give me hard music to learn.

_____85. If my friends chose a music leader, they would probably pick me.

_____86. Other kids sometimes ask me to play or sing with them.

_____87. I expect to play or sing in performing groups in high school/college.

_____88. No one pays much attention to my musical activities at home.

_____89. I don't have even one friend who would say I am any good in music.

_____90. I practice more because my teacher thinks I can do well in music.

_____91. Compared with other kids, I think I am talented.

Appendix F

Advanced Measures of Music Audiation*

*****Adapted from *Advanced Measures of Music Audiation* by E. E. Gordon (1989). Chicago, IL: GIA Publications, Inc.

This is a listening test. You will hear one musical phrase and then a second phrase. You will have three choices with which to compare the two phrases: same, tonally different (pitches), or rhythmically different (rhythms). Listen to the phrases carefully. You will hear each phrase one time.

Hint 1: When you click on the screen to answer, the playback control screen is minimized. You can re-open the playback control screen by clicking it on the taskbar.

Hint 2: You must complete the practice exercises. If you are unsure of an answer during the test itself, it is better to leave the answer blank than guess what the answer might be. Only answer those questions you are sure of.

Introduction (audio link)

Introduction – Click on the word "Introduction above and use this answer sheet.

	Same	Tonal	Rhythm
Example			

Practice 1 (audio link)

Practice Exercise 1 – Click on Practice 1 above and use this answer sheet.

	Same	Tonal	Rhythm
1.			

Practice 2 (audio link)

Practice Exercise 2 – Click on Practice 2 and use this answer sheet.

	Same	Tonal	Rhythm
2.			

Practice 3 (audio link)

Practice Exercise 3 – Click on Practice 3 above and use this answer sheet.

	Same	Tonal	Rhythm
3.			

Music Examples (audio link)

Music Examples 1-10 – Click on Music Examples 1-10 and use this answer sheet.

	Same	Tonal	Rhythm
1.			
2.			
3.			
4.			
5.			
6.			
7.			
8.			
9.			
10.			

Music Examples 11-20 (audio link)

Music Examples 11-20 – Click on Music Examples 11-20 and use this answer sheet.

	Same	Tonal	Rhythm
11.			
12.			
13.			
14.			
15.			
16.			
17.			
18.			
19.			
20.			

Music Examples 21-30 (audio link)

Music Examples 21 -30 – Click on Music Examples 21-30 and use this answer sheet.

	Same	Tonal	Rhythm
21.			
22.			
23.			
24.			
25.			
26.			
27.			
28.			
29.			
30.			

Appendix G

Self-Efficacy Rating Form

Teachers: Please rate your students' self-efficacy in music performance based upon their music behavior in class. Self-efficacy is an individual's belief in his or her ability to be successful in completing a task. For example, a student who volunteers regularly to perform on their instrument would be perceived to have a high level of self-efficacy, while a student who never volunteers would be perceived to have a low level of self-efficacy.

- Have students enter name and birth date when turning in parent permission forms.
- Rate only those students returning permission forms on a 1-5 scale.
  1 = very low, 2 = low, 3 = moderate, 4 = high, 5 = very high
- Cut off students name and return birth date and rating information to researcher. Researcher should not be given student names matched with birth dates.

cut here

Student Name – First, Last	Birth Date (mm/dd/yyyy)	Self-Efficacy Rating

Appendix H

Instructions for Self-Efficacy in Music Performance Study

Hello!
Your participation in this research study will make a valuable contribution to music education.

I. **Package contents**: You will find the following items in your package.

A. Introduction Script
B. Self-Efficacy Rating Forms (for teacher)
C. Parental Permission Forms
D. Assent to Participate Forms (student agreement)
E. Informed Consent to Participate Form (teacher agreement)
F. Check for return postage (only schools outside of the Tampa, FL area)

II. **Directions**: These procedures have been approved by your school district and the Institutional Review Board at the University of South Florida to protect the privacy of your students and the confidentiality of the data. Please follow them.

A. Read the Introduction Script to students
B. Immediately after reading the script, distribute the Parental Permission Forms. You set the deadline for the students to return the forms.
C. Assent to Participate form. Every student returning a signed parent permission form must also complete an assent form to participate.
D. Informed Consent to Participate Form: For teacher
E. Self-Efficacy Rating Form – Use this form to rate your perceptions of your students' self-efficacy beliefs.
F. Allow only those students with signed parent permission and signed assent forms to complete the online questionnaires. Students must work individually. Since one of the questionnaires is a listening test, the students will need headphones.
The web address for the questionnaires is:
http://www.surveygizmo.com/s3/499644/Self-Efficacy-in-Music-Performance
G. Send the (a) the signed parent permission forms, (b) the signed assent forms, (c) your signed informed consent form, and (d) your student self-efficacy rating forms to: Michael Zelenak 5555 Main St., St. Pete Beach, FL 33706. (I will pick up forms from Tampa Bay area schools. Send me an email when you are ready for pick-up.)

*Hint for paperwork collection: When students return parent permission forms, have them sign the assent form, and then they write their name and birth date on the Self-Efficacy Rating form. You can go back and fill in your rating later, cut

off the students' names with scissors, and return the form to me. I will match the students' online questionnaire scores with your ratings using birth dates.

III. **Questions**: You can call me at 727-555-5716 or by email at zelenak@mail.usf.edu.

- Make extra copies if you need them, or let me know and I will provide them.
- If a student gets error messages while completing a questionnaire, it may be because they used the "Enter" key. Disregard the message.
- If students do not complete the questionnaires in one class period, they can click the Save and Continue bar at the top of the page. A window will open asking for an email address. A link will be sent to that address that returns the student to the place where they stopped.

Appendix I

Introduction Script

Hello,

My name is Michael Zelenak and I am a doctoral student at the University of South Florida. I am working on my dissertation project and I need your help. Today, I am asking you to participate in an important project that involves music students from several different schools. Sometimes students feel that no one cares about what they think or feel. This is one opportunity in which your voice can be heard.

- What is this study about?

Experts have found that students who have strong beliefs in their music abilities become better performers than those who do not. I have developed a scale, or test, to help teachers understand their students' beliefs in their musical abilities. In this project, I am testing my scale to see if it works the way it is supposed to work. Even though it is a test, the results will not influence your grade in this class.

- What will you be asked to do?

If you decide to participate, you will log onto a secure web site and provide general information about yourself. This information will include your birthdate, gender, race, grade level, music ensemble, and music teacher's name. You will then complete four questionnaires. Two questionnaires will ask you about your beliefs in your music abilities, one questionnaire will ask you about your beliefs in your mathematics ability, and one questionnaire will evaluate your music listening skills. You will complete these

online questionnaires on computers at school in your regularly scheduled class periods. I will also ask your teacher to provide me with a rating of their perception of your self-efficacy in music performance.

- What else do you need to know?

Your decision to participate is entirely up to you. Here are some things you need to know: (a) participation will not impact your grade, (b) you will answer the questionnaires anonymously (no one will know who you are), (c) I am the only person who will have access to the questionnaire responses, (d) you will not receive any compensation for your participation, and (e) I will obtain approvals from my university, your school, your parents, and you, before the project begins. I may share the class results of this study with your teacher, but no individual results will be shared. I promise!

- Where do we begin?

The first step toward participation is obtaining your parent's permission. Today, your teacher will distribute permission forms that explain the study in detail. Only those students with signed permission forms will be allowed to participate in the study. Please return your form as soon as possible.

This is your opportunity to make an important contribution to music education in our country and around the world. Take that first step: Bring back your parent permission form. Your voice needs to be heard!

Thank-you,

Michael Zelenak
Doctoral Candidate
University of South Florida

Appendix J

Parental Permission to Participate in Research Involving Minimal Risk

Information for parents to consider before allowing their child to take part in this research study

IRB Study # Pro 00003140

The following information is being presented to help you and your child decide whether or not your child wishes to be a part of a research study. Please read this information carefully. If you have any questions or if you do not understand the information, we encourage you to ask the researcher.
We are asking you to allow your child to take part in a research study called: **Self-Efficacy in Music Performance: Measuring the Sources among Secondary School Music Students**
The person in charge of this research study is Michael Zelenak**.** This person is called the Principal Investigator. However, other research staff may be involved and can act on behalf of the person in charge. He is being guided in this research by Dr. C. Victor Fung.

Why is this research being done?

The purpose of this study is to develop a greater understanding of the self-efficacy beliefs in music performance among secondary school students in band, chorus, and string orchestra. Previous research has shown that self-efficacy (i.e., the belief in one's abilities) is the best predictor of achievement in music performance. Students with strong beliefs in their abilities reach higher levels of achievement. This study will evaluate a questionnaire designed to measure the sources of information that contribute to the development of self-efficacy beliefs. With this information, music teachers will be able to identify weaknesses in their students' self-efficacy beliefs and create instruction to address their students' needs.

Why is your child being asked to take part?

We are asking your child to take part in this research study because he or she participates in a band, chorus, or string orchestra program at school.

Should your child take part in this study?

This informed consent form tells you about this research study. You can decide if you want your child to take part in it. This form explains:

- Why this study is being done.
- What will happen during this study and what your child will need to do.
- Whether there is any chance your child might experience potential benefits from being in the study.
- The risks of having problems because your child is in this study.

Before you decide:

- Read this form.
- Have a friend or family member read it.
- Talk about this study with the person in charge of the study or the person explaining the study. You can have someone with you when you talk about the study.
- Talk it over with someone you trust.
- Find out what the study is about.
- You may have questions this form does not answer. You do not have to guess at things you don't understand. If you have questions, ask the person in charge of the study or study staff as you go along. Ask them to explain things in a way you can understand.
- Take your time to think about it.

The decision to provide permission to allow your child to participate in the research study is up to you. If you choose to let your child be in the study, then you should sign this form. If you do not want your child to take part in this study, you should not sign the form.

What will happen during this study?

Your child will be asked to spend about 45 minutes participating in this study. They will complete the survey at school during their regularly scheduled classes using a school computer. The computer may be in the music classroom, a computer lab, or library. Since multiple people use the same computer, your child's identity cannot be traced to the computer. Your child will be asked for his/her birthdate, gender, race, grade, music ensemble, and music teacher's name. This information will be used only to categorize participants into groups. Students will then respond to 4 questionnaires. The questionnaires are:

- Music Performance Self-Efficacy Scale – the results generated by this questionnaire are the primary focus of this study. It asks students to rate past

music experiences, their interactions with other musicians, feedback received from others, and how performing music makes them feel. About 50 students will complete this scale again two weeks later to check the consistency of their scores.

- Self-Esteem of Musical Ability – results from this questionnaire will be compared to those from the Self-Efficacy Scale. The types of questions are very similar to the Self-Efficacy Scale.
- Sources of Mathematics Self-Efficacy Scale – this questionnaire will ask your child about their mathematics experiences and their confidence in their mathematics ability.
- Advanced Measures of Music Audiation – this questionnaire is a listening test. It asks students to identify whether two musical examples are the same or different. If the student thinks the two examples are different, then they must specify if the difference is in pitch or rhythm.

In addition to the questionnaires, music teachers will rate their students on a 1-5 scale depending on their perceptions of students' self-efficacy for music performance. These perceptions are based on the student's behaviors in class. These ratings will only be used for research and not school grades. The researcher may share class results with the teacher, but no individual results will be shared.

How many other people will take part?

Approximately 1,000 students will participate in this study at various schools.

What other choices do you have if you decide not to let your child to take part?

If you decide not to let your child take part in this study, that is okay. There are no consequences for not participating. Your child will participate in their regular school activities.

Will your child be compensated for taking part in this study?

You will receive no payment or other compensation for taking part in this study.

What will it cost you to let your child take part in this study?

It will not cost you anything to let your child take part in the study.

What are the potential benefits to your child if you let him / her take part in this study?

We do not know if your child will gain any benefits by taking part in this study.

What are the risks if your child takes part in this study?

There are no known risks to those who take part in this study.

What happens if you decide not to let your child take part in this study?

You should only let your child take part in this study if both of you want to. You or your child should not feel that there is any pressure to take part in the study to please the study investigator or teacher.

You can get the answers to your questions, concerns, or complaints.

If you have any questions, concerns or complaints about this study, call Michael Zelenak at 727-555-5716 or email him at zelenak@mail.usf.edu. If you have questions about your child's rights, general questions, complaints, or issues with this study, call the USF IRB at (813) 974-5638.

Consent for My Child to Participate in this Research Study

It is up to you to decide whether you want your child to take part in this study. If you want your child to take part, please read the statements below and sign the form if the statements are true.

I freely give my consent to let my child take part in this study. I understand that by signing this form I am agreeing to let my child take part in research. I have received a copy of this form to take with me.

_____ _____

Signature of Parent of Child Taking Part in Study Date

_____

Printed Name of Parent of Child Taking Part in Study

The signature of only one parent was obtained because:

- ☐ The other parent is not reasonable available. Explain:_____
- ☐ The other parent is unknown.
- ☐ The other parent is legally incompetent.
- ☐ The parent who signed has sole legal responsibility for the care and custody of the child.

Appendix K

Assent to Participate in Research

Information for Persons under the Age of 18 Who Are Being Asked To Take Part in Research

IRB Study # Pro00003140

Title of study: Self-Efficacy in Music Performance: Measuring the Sources among Secondary School Music Students

Why am I being asked to take part in this research?

You are being asked to take part in this research study because you participate in a band, chorus, or string orchestra at your school. If you take part in this study, you will be one of about 1,000 students in this study.

Who is doing this study?

The person in charge of this study is Michael Zelenak of the University of South Florida. He is being guided in this research by Dr. C. Victor Fung

What is the purpose of this study?

By doing this study, we hope to learn more about middle and high school music students' beliefs in their musical abilities, sometimes called their music "self-efficacy."

Where is the study going to take place and how long will it last?

The study will be take place at your school during your regular music class period. The total amount of time you will be asked to volunteer for this study is 45 minutes.

What will you be asked to do?

You will log on to a secure web site from a computer at your school. The computer may be in the music classroom, computer lab, or library. Since multiple people use the same computer, your identity cannot be traced to the computer. On the web site, you will be asked for information such as your birthdate, gender, race, grade, music ensemble, and music teacher's name. This information will only be used to categorize participants. You will complete 4 questionnaires. The Music Performance Self-Efficacy Scale will ask you

questions about your beliefs in your musical ability. About 50 students will re-take this scale about two weeks later during the same class period to check for consistency. The Self-Esteem of Musical Ability Questionnaire will ask you how you think about yourself as a musician. The Sources of Self-Efficacy in Mathematics Scale will ask you questions about your belief in your ability to do mathematics. Finally, the Advance Measures of Music Audiation will check your listening skills. Your teacher will also supply ratings of his or her perceptions of your self-efficacy beliefs. The results from all of these sources will be used to compare the beliefs of different groups of students and whether the Music Performance Self-Efficacy Scale is doing a good job in measuring those beliefs. Some class results may be shared with your teacher, but no individual results will be shared.

What things might happen that are not pleasant? To the best of our knowledge, the things you will be doing will not harm you or cause you any additional unpleasant experience. They will be no different than doing a computer assignment in your other classes.

Will something good happen if I take part in this study? We cannot promise you that anything good will happen if you decide to take part in this study.

What other choices do I have if I do not participate?

You may choose not to participate in this research study. You will continue with your school work.

Do I have to take part in this study?

You should talk with your parents or anyone else that you trust about taking part in this study. If you do not want to take part in the study, that is your decision. You should take part in this study because you really want to volunteer.

If I don't want to take part in this study, what will happen?

If you do not want to be in the study, nothing else will happen.

Will I receive any rewards for taking part in this study?]

You will not receive any reward for taking part in this study.

Who will see the information about me?

Your information will be added to the information from other people taking part in the study so no one will know who you are.

Can I change my mind and quit?

If you decide to take part in the study you still have the right to change your mind later. No one will think badly of you if you decide to quit.

What if I have questions?

You can ask questions about this study at any time. Before participating in the study, talk with your parents, teachers, or other adults that you trust. You can contact the principal investigator of the study Michael Zelenak at 727-555-5716. If you think of questions during the study, you can ask your teacher.

Assent to Participate

I understand what the person running this study is asking me to do. I have thought about this and agree to take part in this study

_____	_____
Name of person agreeing to take part in the study	Date
_____	_____
Name of person providing information to subject	Date

Appendix L

Informed Consent to Participate in Research
Information to Consider Before Taking Part in this Research Study

IRB Study # Pro00003140

You are being asked to take part in a research study. Research studies include only people who choose to take part. This document is called an informed consent form. Please read this information carefully and take your time making your decision. Ask the researcher or study staff to discuss this consent form with you, please ask him/her to explain any words or information you do not clearly understand. We encourage you to talk with your family and friends before you decide to take part in this research study. The nature of the study, risks, inconveniences, discomforts, and other important information about the study are listed below.

We are asking you to take part in a research study called: **Self-Efficacy in Music Performance: Measuring the Sources among Secondary School Music Students**

The person who is in charge of this research study is Michael Zelenak**.** This person is called the Principal Investigator. However, other research staff may be involved and can act on behalf of the person in charge. He is being guided in this research by Dr. C. Victor Fung. The research will be conducted at your school.

Purpose of the study

The purpose of this study is to:

- Develop a great understanding of self-efficacy beliefs in music performance among secondary school music students.
- You are being asked to participate in this study because you are a teacher of secondary school music students.
- This study is being conducted by a student as his doctoral dissertation.

Study Procedures

If you take part in this study, you will be asked to:

- Read an introductory script to your students
- Distribute and collect parent permission and student assent forms
- Provide computer access for students to complete online questionnaires.
- Provide ratings on a 1-5 scale of your students' self-efficacy beliefs in music performance. The researcher will provide you with a form to collect this information. You will identify students by their birthdates and indicate their rating. You can complete this form at your convenience and return it as an email attachment to the researcher. You will only rate students who have turned in signed parent permission forms. Depending on the return rate of permission forms, your time commitment could range from 15 to 60 minutes.

Total Number of Participants

About 1,000 students will take part in this study from various schools in the United States.

Alternatives

You do not have to participate in this research study.

Benefits

We are unsure if you will receive any benefits by taking part in this research study.

Risks or Discomfort

This research is considered to be minimal risk. That means that the risks associated with this study are the same as what you face every day. There are no known additional risks to those who take part in this study.

Compensation

You will receive no payment or other compensation for taking part in this study.

Cost

There will be no cost to you as a result of being in this study.

Privacy and Confidentiality

We will keep your study records private and confidential. Certain people may need to see your study records. By law, anyone who looks at your records must keep them completely confidential. The only people who will be allowed to see these records are:

- The Principal Investigator and his advisor.
- Certain government and university people who need to know more about the study. For example, individuals who provide oversight on this study may need to

look at your records. This is done to make sure that we are doing the study in the right way. They also need to make sure that we are protecting your rights and your safety.

- Any agency of the federal, state, or local government that regulates this research. This includes the Department of Health and Human Services (DHHS) and the Office for Human Research Protection (OHRP).
- The USF Institutional Review Board (IRB) and its related staff who have oversight responsibilities for this study, staff in the USF Office of Research and Innovation, USF Division of Research Integrity and Compliance, and other USF offices who oversee this research.

We may publish what we learn from this study. If we do, we will not include your name. We will not publish anything that would let people know who you are.

Voluntary Participation / Withdrawal

You should only take part in this study if you want to volunteer. You should not feel that there is any pressure to take part in the study. You are free to participate in this research or withdraw at any time. There will be no penalty or loss of benefits you are entitled to receive if you stop taking part in this study.

New information about the study

During the course of this study, we may find more information that could be important to you. This includes information that, once learned, might cause you to change your mind about being in the study. We will notify you as soon as possible if such information becomes available.

You can get the answers to your questions, concerns, or complaints

If you have any questions or concerns about this study call Michael Zelenak at 727-555-5716.

If you have questions about your rights as a participant in this study, general questions, or have complaints, concerns or issues you want to discuss with someone outside the research, call the USF IRB at (813) 974-5638.

Consent to Take Part in this Research Study

I freely give my consent to take part in this study. I understand that by signing this form I am agreeing to take part in research. I have received a copy of this form to take with me.

_____ _____

Signature of Person Taking Part in Study Date

_____

Printed Name of Person Taking Part in Study

Appendix M

Using Excel to Calculate Self-Efficacy Scores

Background information: Columns contain item results and are identified by letter names. Rows contain student results and are identified by numbers. This example is for a class of 10 students. Adjust the class size as necessary.

I. Label cells A1-X1 as Item 1 - Item 24.

II. Enter each student's data as rows, 2-11.

III. Calculate column totals. In cell A12, enter "= SUM(A2:A11)" for column A, in cell B12, enter "=SUM(B2:B11)" for column B, etc. through column X.

IV. Calculate item averages. In cell A13, enter"= AVERAGE(A2:A11)" for column A, in cell B13 enter"=AVERAGE(B2:B11)" for column B, etc. through column X.

V. Calculate sources of self-efficacy totals.

A. Create labels. In cell A14, enter "Mastery", in cell A15, enter "Vicarious", in cell A16 enter "Verbal", in cell A17, enter "Phys", in cell A18, enter "Total".

B. Calculate sources of self-efficacy totals.

1. Mastery experience: In cell B14, enter "= SUM(A12, D12, F12, H12, J12, L12, N12, P12)".
2. Vicarious Experience: In cell B15, enter "= SUM(B12, E12, K12, R12, T12)".
3. Verbal/Social Persuasion: In cell B16, enter "= SUM(C12, G12, I12, M12, U12, V12)".
4. Physiological State: In cell B17, enter "= SUM(O12, Q12, S12, W12, X12)".
5. Total: In cell B18, enter "= SUM(A12-X12)".

C. Calculate sources of self-efficacy mean scores.

1. Mastery Experience: In cell C14 enter "= B14/80". (The denominator is the number of students times the number of items – 10 students * 8 items)
2. Vicarious Experience: In cell C15 enter "= B15/50)"
3. Verbal/Social Persuasion: In cell C16, enter "= B16/60)".
4. Physiological State: In cell C17, enter "= B17/50)".
5. Total: In cell C18, enter "= B18/240".

VI. Calculate student scores.

A. Create labels. In Y1, enter “Mastery”, in Z1, enter “Vicarious”, in AA1, enter “Verbal”, in AB1, enter “Phys”, and in AC1, enter “Total”

B. Calculate student mean scores.

1. Mastery Experience: For the student whose scores are listed in row 2, enter “= SUM(A2, D2, F2, H2, J2, L2, N2, P2)/8” into cell Y2. Change row numbers for each student.
2. Vicarious Experience: For the student whose scores are listed in row 2, enter “= SUM(B2, E2, K2, R2, T2)/5” into cell Z2.
3. Verbal/Social Persuasion: For the student whose scores are listed in row 2, enter “= SUM(C2, G2, I2, M2, U2, V12)/6” into cell AA2.
4. Physiological State: For the student whose scores are listed in row 2, enter “= SUM(O2, Q2, S2, W2, X)/5” into cell AB2.
5. Total: For the student whose scores are listed in row 2, enter “= SUM(A2-X2)/24” into cell AC2.

VII. Compare student scores (columns Y-AC) to class mean scores (cells C14-C18) or use mean scores in this study (Table 20) to identify strengths and weaknesses.

Appendix N

Inter-Item Correlation Coefficients

Inter-Item Correlation Matrix for the MPSES (N = 290)

	I1	I2	I3	I4	I5	I6	I7	I8	I9
I1	-								
I2	.31	-							
I3	.36	.31	-						
I4	.44	.13	.16	-					
I5	.22	.37	.24	.22	-				
I6	.34	.35	.47	.14	.29	-			
I7	.33	.16	.40	.20	.19	.33	-		
I8	.28	.11	.14	.36	.21	.19	.29	-	
I9	.32	.33	.42	.29	.27	.34	.31	.29	-
I10	.48	.20	.41	.28	.24	.40	.26	.22	.31
I11	.13	.17	.11	.21	.26	.09	.06	.06	.17
I12	.36	.36	.32	.32	.24	.23	.32	.28	.49
I13	.20	.10	.40	.18	.16	.34	.43	.23	.24
I14	.22	.29	.29	.15	.26	.27	.10	.05	.34
I15	.19	.13	.14	.20	.17	.18	.17	.06	.12
I16	.32	.35	.40	.17	.25	.46	.29	.14	.32
I17	.34	.25	.40	.27	.19	.35	.39	.25	.32
I18	.17	.19	.23	.11	.17	.19	.14	.12	.17
I19	.13	.09	.21	.10	.12	.17	.21	.10	.12
I20	.13	.15	.23	.14	.11	.17	.04	.10	.17
I21	.33	.16	.39	.18	.21	.34	.34	.18	.34
I22	.17	.23	.43	.10	.20	.35	.37	.08	.29
I23	.47	.27	.42	.32	.27	.34	.38	.29	.31
I24	.37	.12	.28	.34	.28	.28	.38	.24	.29

Inter-item Correlation Matrix for the MPSES (N = 290)

	I10	I11	I12	I13	I14	I15	I16	I17	I18
I10	-								
I11	.22	-							
I12	.36	.19	-						
I13	.33	.15	.25	-					
I14	.25	.19	.32	.21	-				
I15	.33	.14	.27	.18	.18	-			
I16	.30	.17	.31	.29	.31	.22	-		
I17	.36	.16	.45	.34	.27	.24	.42	-	
I18	.16	.22	.21	.13	.16	-.04	.24	.31	-
I19	.08	.04	.14	.19	.17	.34	.13	.12	.00
I20	.16	.33	.16	.06	.21	.06	.13	.20	.26
I21	.33	.18	.28	.35	.28	.20	.26	.44	.24
I22	.23	.11	.24	.37	.25	.19	.38	.38	.26
I23	.41	.12	.37	.39	.20	.25	.30	.66	.19
I24	.31	.11	.26	.35	.25	.25	.21	.51	.17

Inter-item Correlation Matrix for the MPSES (N = 290)

	I19	I20	I21	I22	I23	I24
I19	-					
I20	.10	-				
I21	.09	.25	-			
I22	.21	.22	.49	-		
I23	.13	.17	.48	.38	-	
I24	.17	.03	.43	.36	.59	-

Appendix O

Means, Standard Deviations, and Cell Sizes for MPSES

Means, Standard Deviations, and Sample Sizes for 3x2 Multivariate Analysis of Variance for Music Performance Self-Efficacy Scale (N = 290)

Grade	Band			Chorus			Strings			Marginal		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Middle	7			122			21			150		
Mastery		514	79		604	145		657	106		607	140
Vicarious		267	87		346	100		346	100		349	98
Verbal/Social		407	106		477	111		508	76		478	107
Physiological		339	120		393	120		406	82		392	98
High	80			31			29			140		
Mastery		623	87		627	129		620	107		623	101
Vicarious		375	78		386	71		363	101		375	82
Verbal/Social		480	92		502	109		473	90		483	96
Physiological		390	76		400	68		419	78		398	75
Marginal	87			153			50			290		
Mastery		614	91		609	142		636	107		615	123
Vicarious		367	84		354	96		375	89		361	91
Verbal/Social		474	95		482	110		488	85		481	102
Physiological		385	81		394	94		414	79		395	88

About the Author

Michael S. Zelenak was born in Detroit, Michigan. He earned a B.M. degree in Music History/Musicology from the University of Michigan, an M.A. degree in Music Education from the University of South Florida, and a Ph.D. in Music from the University of South Florida. He taught at Dunedin Highland Middle School from 2002-2008 and was certified by the National Board for Professional Teaching Standards. As a graduate assistant, he was awarded the Distinguished Graduate Achievement Award from the USF College of The Arts and the Outstanding Merit Award in Music Education from the USF School of Music. Prior to entering music education, he pursued a career as a performer, composer, and arranger of popular and jazz music styles. Along with self-efficacy, his current interests include research methodologies, music technology, and preparing pre-service teachers to enter the field of music education.