Development and Validation of the Music Performance Self-Efficacy Scale

Michael S. Zelenak
Dixie M. Hollins High School—St. Petersburg, FL, U.S.A.

Abstract

*The purpose of this study was to develop and validate the Music Performance Self-Efficacy Scale (MPSES). The MPSES was designed to measure Bandura's (1977, 1986) four sources of self-efficacy (i.e., mastery experiences, vicarious experiences, verbal/social persuasion, and physiological state) in music performance among secondary school students. Participants (*N *= 293) attended a middle school in the southeastern United States. Scores generated from the MPSES were internally consistent (α = .97) and indicated strong relationships between the sources and the composite construct. Results of confirmatory factor analysis provided evidence of good fit between the data and Bandura's model* χ^2 *(224) = 568.49,* p *< .001, CFI = .95, RMSEA = .07, and SRMR = .04. This may be the most important finding of the study. It confirmed that Bandura's four sources of self-efficacy contribute to the development of self-efficacy for music in ways similar to other subjects. Additional evidence supporting the use of the MPSES as a measure of music performance self-efficacy included (a) significant differences between music and non-music students on measures of the sources of self-efficacy,* Λ *= 0.63,* F*(4,287) = 42.88,* p *< .001, (b) a positive correlation between teacher ratings of student self-efficacy and MPSES scores (*r *= .44), and (c) discriminant evidence established by comparisons of correlations between MPSES scores and scores from other self-efficacy measures. These findings suggest that the MPSES is a valid and reliable measure of self-efficacy in music performance among secondary school students. Evaluation with a larger and more diverse sample was recommended.*

Secondary school students participate in a variety of music ensembles. These ensembles include band, chorus, and string orchestra. The instruction the students receive in these ensembles, however, often focuses on the development of psychomotor skills while neglecting the students' self-perceptions of their abilities (Schmitt, 1979). Students spend hours refining their physical coordination, aural acuity, and visual discernment with little time spent nurturing their beliefs in those abilities. Recent studies have shown that self-perceptions are an important part of learning in music performance (Hendricks, 2009; McCormick & McPherson, 2003; McPherson & McCormick, 2000, 2006; Wehr-Flowers, 2007). In particular, the self-perception of self-efficacy as described by Bandura (1986) has been linked to achievement in music performance. McPherson and McCormick (2000) found performance self-efficacy as the best predictor of music achievement among music students ages 9 to 18 years. They stated, "How students think about themselves, the task, and their performance is just as important as the time they devote to practicing their instrument" (p. 31). Although other researchers have made similar comments (Greenberg, 1970; Schmitt, 1979; Svengalis, 1978), it is surprising that instruction related to self-efficacy has not received greater attention in the music curriculum. One explanation may be the absence of a valid and reliable measure of music performance self-efficacy. A scale designed specifically to measure this construct would be a useful tool for educators and researchers. It would provide insight into student self-perceptions and assist educators in improving student achievement. In this study, therefore, I have documented the development and validation of a scale I constructed to measure self-efficacy in music performance—the Music Performance Self-Efficacy Scale (MPSES).

Bandura's (1986) social cognitive theory was used as the framework for the Music Performance Self-Efficacy Scale (MPSES). This decision was based on several rationales. First, scales based on this theory in other subject areas have been found to produce valid and reliable results (Aydin & Uzuntiryaki, 2009; Lent, Lopez, Brown, & Gore, 1996; Pajares, 2007; Usher & Pajares, 2006, 2009). Second, previous findings in music education research have been consistent with the tenets of this theory (Craske & Craig, 1984; Hendricks, 2009; McCormick & McPherson, 2003; McPherson & McCormick, 2000, 2007). And third, this theory resonated with my personal experiences.

In his theory, Bandura (1986) proposed a reciprocal model in which cognition, behavior, and environment, interacted to influence human thought and action. Self-efficacy was self-referent thought that mediated the relationship between cognition and behavior. Bandura defined 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 practice, individuals make behavioral decisions based on their beliefs in their abilities to accomplish specific tasks.

Bandura (1977, 1986) identified four sources of information that contributed to the development of

self-efficacy beliefs. Those sources were mastery experiences, vicarious experiences, verbal/social persuasion, and physiological state. The MPSES was divided into four distinct sections with each section reflecting one of the four sources: items 1-7 reflected mastery experiences, items 8-12 reflected vicarious experiences, items 13-18 reflected verbal/social persuasion, and items 20-24 reflected physiological state. Bandura believed mastery experiences had the strongest influence on self-efficacy beliefs and were based on one's prior successes or failures. Vicarious experiences were less influential and consisted of the predictions of success or failure individuals make for themselves based on their observations of others similar to themselves engaged in specific activities. Verbal/social persuasions referred to the judgments and opinions of others and their influence on an individual's decision-making process. And, physiological states were the feelings individuals experienced when engaged in, or thinking about, an activity or behavior.

Bandura (2006) offered a set of guidelines for the development of self-efficacy scales. He characterized self-efficacy as a "set of self-beliefs linked to distinct realms of functioning" (p. 307) rather than a global trait. These beliefs were tied to specific knowledge domains and physical activities. He also believed self-efficacy had a unique identity and that scales of self-efficacy should differentiate it from similar constructs such as self-esteem, confidence, and outcome expectancies. Bandura recommended that self-efficacy scales included items covering a range of generality, strength, and level. Generality referred to variations in the functioning associated with an activity, strength addressed the certainty of an individual to accomplish the task, and level captured the inherent difficulty of the task. He suggested using an item response format that allowed individuals to rate themselves in 10 unit intervals over a 0-100 range (pp. 312-313). In evaluating a scale, he believed that it was important to verify the homogeneity of the items, analyze the items' ability to differentiate between participants, and determine that items reflecting the same domain correlated with each other and with the total score.

Researchers have constructed self-efficacy scales to measure a variety of content areas and activities. Initially, these scales were developed for use in psychology. Sherer and others (1982) created one of the first measures of self-efficacy. It was intended for use among therapists to provide an index of their clients' therapeutic progress. Bandura developed two general self-efficacy scales. His Children's Perceived Self-Efficacy scales (1990, as cited in Pastorelli, Caprara, Barbaranelli, Rola, Rozsa, & Bandura, 2001) was used to measure self-efficacy in preadolescent children and his Multidimensional Scales of Perceived Self-Efficacy (1990, as cited in Choi, Fuqua, & Griffin, 2001) was intended for use with older children and younger adults. Self-efficacy scales were also designed for use in education. These scales measured academic self-efficacy (Pintrich, Smith, Garcia, & McKeachie, 1991), chemistry self-efficacy (Aydin & Uzuntiryaki, 2009; Dalgety, Coll, & Jones, 2003), mathematics self-efficacy (Betz & Hackett, 1983), and writing self-efficacy (Pajares & Valiante, 1999).

Educational researchers have constructed scales to measure Bandura's four sources of self-efficacy within specific subject areas. They have examined the sources of self-efficacy in mathematics (Lent, Lopez, & Bieschke, 1991; Lopez & Lent, 1992; Matsui, Matsui, & Ohnishi, 1990; Usher & Pajares, 2009) and in academics (Usher & Pajares, 2006). Findings from these studies have fell within Banudura's proposed guidelines. Results from Usher and Pajares' (2006) Sources of Self-Efficacy Scale exhibited significant ($p < .001$) correlations between the four sources of self-efficacy and the composite construct (mastery $r = .57$, vicarious, $r = .39$, verbal $r = .45$, physiological $r = -.39$). In another study, data generated by Usher and Pajares' (2009) Sources of Middle School Mathematics Self-Efficacy Scale demonstrated good fit with Bandura's proposed model, $\chi^2(246) = 601.21$, $p < .0001$, CFI = .96, SRMR = .04, and RMSEA =.04. These fit index values met Hu and Bentler's (1998) benchmarks, CFI $\geq$ .95, SRMR $\leq$ .08, and RMSEA $\leq$ .06.

Self-efficacy has been related to achievement in a variety of subject areas. It was correlated with reading grade, $r = .32$, $p < .001$ (Usher & Pajares, 2006), mathematics grade, $r = .29$, $p < .05$ (Lent & Lopez, 1992), and science achievement test scores, $r = .46$, $p < .001$ (Joo, Bong, & Choi, 2000). It has also been associated with achievement in music education. Initially, Greenberg (1970) attributed low achievement in vocal performance among young boys to low self-concepts of their ability to sing. As noted above, McPherson and McCormick (2000) found performance self-efficacy as the best predictor of achievement on a music performance test. They examined the relationships of anxiety/confidence, intrinsic value, general music self-efficacy, and performance self-efficacy with music performance test results. Through stepwise regression analysis, 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%.

McCormick and McPherson (2003) replicated these results and extended the methodology used in their investigation from regression to structural equation modeling (SEM). Using SEM, they explored

the direction and strength of influences between self-efficacy and variables such as practice time, anxiety, and grade level. McPherson and McCormick (2006) identified one configuration of variables in which the data fit the proposed model, χ^2 (364, N = 686) = 1837.78, $p < .01$, AGFI = .93, and RMSEA = .08. In this model, self-efficacy mediated the influence of formal practice, informal practice, practice regulation, and grade level on the outcome variable of music performance. Self-efficacy beliefs determined, in part, the level of influence each variable had on performance achievement.

Although these studies have made important contributions to our understanding of self-efficacy, one area of concern has been the diversity of data collection techniques. Some researchers have adapted measures from other content areas. Nielsen (2004) altered the academic self-efficacy section from the Motivated Strategies for Learning Questionnaire (Pintrich et. al, 1991), and Ritchie and Williamon (2007) modified the general self-efficacy subscale from Sherer and others' (1982) Self-Efficacy Scale. Other researchers have developed their own questions to investigate self-efficacy in music performance. Initially, McPherson and McCormick (2000) used the question "What result do you think you will get for your exam today" (p. 34) to measure self-efficacy beliefs. They expanded their line of inquiry by adding an additional question (McCormick & McPherson, 2003) and then added several more (McPherson & McCormick, 2006). Wehr-Flowers (2007) combined these two approaches by adapting existing measures and creating her own questions. Although variety in data collection strategies allowed for triangulation of findings, it also made comparing findings across studies difficult. The development of a valid and reliable instrument to measure self-efficacy in music performance would facilitate the collection and comparison of data across studies. The purpose of this study, therefore, was to develop and validate a scale to measure music performance self-efficacy among secondary school music students. This scale was designed to be used by researchers to further their investigations of self-efficacy and for use by educators to assist them in meeting the needs of their students.

Method

Participants

Participants (N = 293) attended a public middle school in the southeastern United States. The school was located in a suburban neighborhood of moderate income families with an overall enrollment of 1,228 students. The participants constituted a convenience sample. Seven teachers, including two music teachers, allowed me to enter their classrooms to administer the questionnaires. Participants were enrolled in grade 6 (n = 165), grade 7 (n = 52), and grade 8 (n = 76) with ages ranging from 11 to 14 years. For the purpose of this study, those participants enrolled in music classes were considered music students (n = 154) and participants not enrolled in music classes were non-music students (n = 139). Music students participated in the following classes: beginning band (n = 85), intermediate band (n = 10), advanced band (n = 27), chorus (n = 8), percussion (n = 13), and steel drum (n = 11). Students enrolled in multiple music classes responded to only one set of questionnaires. The participants were diverse in sex (female n = 165, male n = 128,) and race (Asian n = 5, Black n = 5, Hispanic n = 14, mixed race n = 14, Native American n = 5, Other n = 7, and White n = 231). Some students had difficulty defining their racial identity. The racial proportions of the participants were similar to the racial proportions of the school's population.

Measures

The participants completed three scales: (a) Music Performance Self-Efficacy Scale (see Appendix A), (b) Children's Perceived Self-Efficacy Scales, and (c) Writing Self-Efficacy Scale. The Music Performance Self-Efficacy Scale (MPSES) was the primary focus of this study. Bandura's (1990) Children's Perceived Self-Efficacy Scales (CPSES) were used to measure academic self-efficacy, and Pajares and Valiante's (1999) Writing Self-Efficacy Scale (WSES) was used to measure writing self-efficacy. In addition, the two music teachers provided a list of students they perceived as exhibiting particularly high or low levels of self-efficacy in music performance.

Music Performance Self-Efficacy Scale. The MPSES was created to measure Bandura's four sources of self-efficacy within the context of music performance. Initially, a pool of 30 items was assembled from a variety of resources including a general self-efficacy scale (Sherer et. al, 1982), an academic self-efficacy scale (Pintrich et. al, 1991), a scale on sources of mathematics self-efficacy (Lent, Lopez, & Bieschke, 1991), an academic self-efficacy scale (Usher & Pajares, 2006), and journal articles investigating music performance self-efficacy (McCormick & McPherson, 2003; McPherson & McCormick, 2006). These items were then modified to be diverse in generality, strength, and level, as well as reflect the domain of music performance. This modified pool of 30 items was presented to a panel of five music education professors and five doctoral students in music education to establish evidence of content validity. The revised items were then sent to another professor with expertise in music self-efficacy for further review. The revised scale 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 check the participant's accuracy in responding to the items (#19). This item asked participants to simply write the number "9" as the response. The following item (#20) included directions to return to the original response format.

Participants responded to each item by writing a number between 0-100 on a line before each item. Although Bandura (2006) recommended using a format of 0-100 in 10 unit intervals, several sources provided evidence that convinced me to deviate from his recommendation. First, Usher and Pajares (2009) found the 0-100 response format resulted in a higher level of internal consistency (Mathematics Skills Self-Efficacy, α = .95) than the 6-point Likert-type format (Sources of Mathematics Self-Efficacy, α = .91). Second, Pajares, Hartley, and Valiante (2001) found the 0-100 response format psychometrically stronger than a 6-point Likert-type format based on its ability to account for a larger percentage of explained variance. They confirmed the similarity in middle school students' ability to make discriminating judgments by comparing student scores using the 0-100 format (M = 75.07) and the 6-point Likert-type scale (M = 4.4, or 73.17 on the 0-100 scale). Finally, Byrne (2005) advised against using ordinal level data for confirmatory factor analysis and advocated for the use of interval level data.

Children's Perceived Self-Efficacy Scales. Bandura's CPSES (1990) consisted of 37 items intended to measure 7 domains of self-efficacy. Those domains were: (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. Using exploratory factor analysis, Pastorelli and others (2001) found the 37 items loaded on three underlying factors among an international group of participants 10-15 years of age from Hungary, Italy, and Poland. The authors labeled the factors academic self-efficacy, social self-efficacy, and self-regulatory self-efficacy. The scree test identified three factors with eigenvalues > 1 (Cattell, 1966) and item loadings confirmed the presence of the same three principal factors for the three countries. The researchers conducted Principal Component Analysis (PCA) and Simultaneous Component Analysis (SCA) on the same data set and found the data yielded the same results. In the current study, participants completed all items on the CPSES, but only the 19 items reflecting academic self-efficacy were used in the analysis. Those items included 1-17, 27, and 28. Pastorelli and others (2001) found the responses from these items to have high levels of internal consistency (α = .89 in Poland, α = .87 in Italy, and α = .86 in Hungary). Their values were above the .80 benchmark for acceptable consistency recommended by Henson (2001) for general research applications. These same items were used to measure academic self-efficacy in a previous validation study of a self-efficacy scale (Usher & Pajares, 2006).

Writing Self-Efficacy Scale. Pajares and Valiante's (1999) WSES consisted of 10 items. The items measured students' self-perceptions of their grammatical and compositional skills in writing. Participants responded to items using a response format of 0 (no chance) to 100 (completely certain). In an exploratory factor analysis, Pajares (2007) found that the scores from the 10 items loaded on 2 factors with eigenvalues of 13.0 and 1.9 which accounted for 99% of the variance among students in grades 4 through 11. Factor 1 (items 1-5) reflected grammar skills and Factor 2 (items 6-10) reflected compositional skills. The internal consistency of the responses for the total scale was α = .91, for Factor 1 α = .88, and for Factor 2 α = .86.

Teacher ratings. As an alternate method of data collection, the two music teachers in the current study identified students they perceived as having very high or very low self-efficacy beliefs based upon the student's music behavior in class. The teachers rated students based on behaviors such as volunteering to perform on their instrument as high and students who never volunteer as low. The teachers identified 15% of the participants (n = 43) as having high or low self-efficacy beliefs. These ratings were correlated with student scores on the Music Performance Self-Efficacy Scale.

Procedure

The study took place in the middle of October during the fall semester. The school secretary set up a schedule in which I visited two classrooms per period over three days (Wednesday, Thursday, and Friday). I returned the following week on the same days to administer the scales. In the first meeting, I introduced myself, the goals of the study, and distributed parent consent forms. All students were encouraged to participate. In the second meeting, I collected the parent consent forms, and then had only those students with signed parent forms complete the student assent forms and the scales. The scales were stapled together in all possible combinations to avoid confounding variables related to test order. The students wrote their birth date on the top page allowing me to match the written scales to the ratings from the music teachers. Gall, Gall, and Borg (2007) proposed that children who return parental permission forms differ from those who do not. Given these limitations, the results of this study must be interpreted with caution.

Analysis

Several types of analyses were conducted. The objective of these analyses was to provide evidence in determining whether or not the intended interpretations of scores from the Music Performance Self-Efficacy Scale were valid and reliable. Scores from the MPSES were not intended to be used as measures of achievement but rather as evidence from which to identify strengths and weaknesses in student self-perceptions of self-efficacy.

The results were analyzed to identify missing data, recognize outliers, determine the normality of the data distribution, and examine the relationships among the items. Confirmatory factor analysis (CFA) was then used to test the fit of the data to the proposed model. CFA was appropriate for this study because the proposed model was based on Bandura's (1986) existing framework. The CFA process consisted of specifying the model (i.e., constructing a visual representation of latent factors, variables, and errors) (see Figure 1), identifying the model (i.e., setting one loading for each factor to 1.0), estimating the model (maximum likelihood estimation was used to estimate the loadings), examining the fit of the data to the proposed model (i.e., comparing fit indices with benchmarks), and making modifications to the model when necessary. CFA was used to determine the loadings between the items and the sources of self-efficacy (1st order), as well as the sources of self-efficacy and the composite construct of self-efficacy (2nd order). The CFA was then extended to evaluate whether or not the items on the MPSES were invariant (i.e., measuring music and non-music students without bias).

The discriminant capabilities of the MPSES were also examined. Using multivariate analysis of variance (MANOVA), the data were evaluated to determine if scores on the MPSES were different between music and non-music students. MANOVA has advantages over other analytical techniques in its ability to compare group means on multiple dependent variables while controlling Type I error. In this case, the dependent variables were the scores reflecting the four sources of self-efficacy. Correlations between scale results were also compared. The relationship between Bandura's CPSES (academic self-efficacy) and Pajares and Valiante's WSES (writing self-efficacy) was compared with the relationship between Bandura's CPSES (academic self-efficacy) and this study's MPSES (music self-efficacy). This comparison provided evidence of validity based on the relationship between different variables. Evidence of convergent validity was established by comparing teacher ratings with student scores on the MPSES. Analytical procedures were conducted using Statistical Analysis Software (SAS), PASW (formerly SPSS), and Mplus.

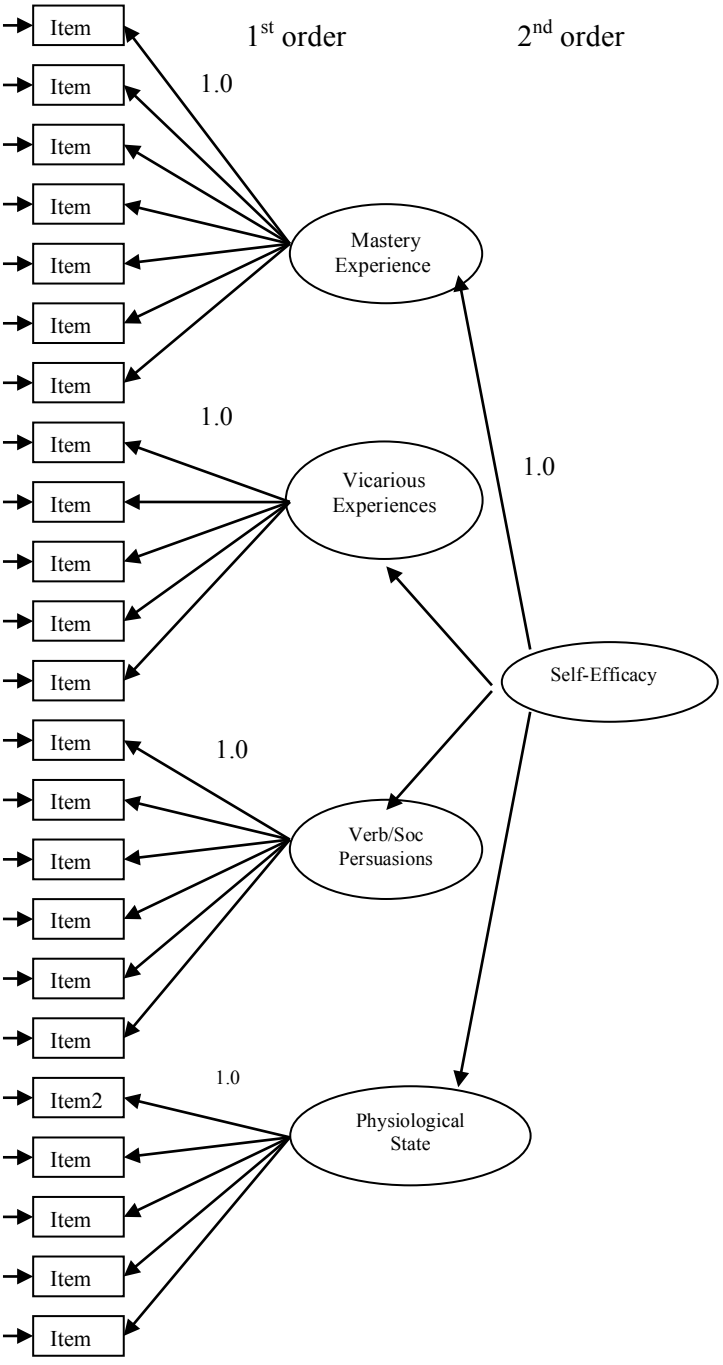

Figure 1. Model specification and identification.

Results

I evaluated the data to determine if the assumptions underlying the analytical techniques were met. All responses were included in the analysis

regardless of missing data. No attempt was made to impute the missing scores. Extreme scores in which participants rated themselves very high or very low were not considered outliers due to the nature of the content and were included in the analyses. At first, the normality of the data distributions came into question. The results from the MPSES, CPSES, and WSES indicated non-normal distributions for each scale based on visual inspection of stem-and-leaf plots, box-and-whisker plots, and the Shapiro-Wilks test for normality ($p < .0001$). In contrast to these results, the skewness and kurtosis values were in acceptable ranges (see Table 1). Stevens (2009) advised that multivariate analyses were robust to violations of non-normality in determining Type I error, but platykurtic kurtosis characteristics (i.e., negative kurtosis values) attenuated the power associated with making Type I error decisions. Therefore, power was attenuated in this study due to this condition. The homogeneity of variance of the data was evaluated using the Brown-Forsythe procedure and was found to be heterogeneous rather than homogenous. Stevens (2009) stated that multivariate tests were robust to violation of this assumption if the samples were similar in size. In this study, the ratio of music to non-music students (1.1) was within Stevens' acceptable range (large group / small group < 1.5). This similarity of sample sizes and the large number of participants supported my decision to proceed with the analytical procedures.

Table 1
Descriptive Statistics of Self-Efficacy, Sources of Self-Efficacy, Academic Self-Efficacy, and Writing Self-Efficacy

Scale	Music Students (n = 154)	Non-Music Students (n = 139)	Total Participants (N = 293)		
	M (SD)	M (SD)	M (SD)	γ_1	γ_2
Music Self-Efficacy (23 items)	1798 (394)	945 (748)	1393 (726)	-0.72	-0.93
Mastery Experience (7 items)	568 (133)	323 (240)	451 (227)	-0.81	-0.68
Vicarious Experience (5 items)	316 (131)	164 (166)	244 (167)	-0.18	-1.36
Verbal Persuasion (6 items)	502 (116)	251 (221)	383 (214)	-0.75	-0.98
Physiological State (5 items)	412 (94)	209 (171)	315 (169)	-0.71	-0.93
Academic Self-Efficacy (19 items)	1471 (272)	1455 (287)	1463 (279)	-0.84	0.56
Writing Self-Efficacy (10 items)	807 (177)	782 (151)	795 (166)	-1.11	1.07

γ_1 = Skewness
γ_2 = Kurtosis

Analysis of individual items on the MPSES indicated strong internal relationships. Internal consistency of the items was high within each section (mastery experience $\alpha = .93$, vicarious experience $\alpha = .90$, verbal/social persuasion $\alpha = .94$, and physiological state $\alpha = .90$) and within the total scale ($\alpha = .97$). Inter-item correlations ranged from $r = .41$ between items 8 and 23, to $r = .83$ between items 13 and 20. Item-to-total correlations ranged from $r = .63$ for item 23, to $r = .89$ for item 17. Correlations between the sources ranged from $r = .75$ between vicarious experiences and physiological state to $r = .89$ between verbal/social persuasion and physiological state (see Table 2). Correlations between the sources and the composite construct ranged from $r = .88$ for vicarious experiences to $r = .96$ for verbal/social persuasion.

Confirmatory factor analysis was conducted using the Mplus software. Mplus automatically identified the model by setting the loading to 1.0 from each factor to one item. The results of the first analysis were χ^2 (226, N = 293) = 650.77, $p < .001$, CFI = .93, RMSEA = .08, and SRMR = .04. In studies with large

samples, χ^2 values can indicate misfit while the other indices suggest good fit. Fit indices accounted for the differences in variance between the data and the proposed model. Hu and Bentler (1998) proposed benchmarks of CFI = .95, RMSEA = .06, and SRMR = .08 to identify good fit. Although values from this study were close, the modification indices indicated strong correlations between items 9 and 10, and items 13 and 19. I allowed these items to correlate, rather than removing any of the correlated items, and the fit improved to more acceptable levels, χ^2 (224, N = 293) = 568.49, p < 0.001, CFI = .95, RMSEA = .07, and SRMR = .04.

Table 2
Correlations between sections and composite of self-efficacy from the MPSES

	Self-Efficacy	Mastery	Vicarious	Verbal/Social	Physiological
Self-Efficacy	1.00				
Mastery	.95	1.00			
Vicarious	.88	.77	1.00		
Verbal/Social	.96	.87	.80	1.00	
Physiological	.94	.86	.75	.89	1.00

All correlations were significant (p < .001).

Factor invariance testing was used to determine if the items were functioning without bias between the music and non-music students. The χ^2 values and degrees of freedom from a CFA for music students, χ^2 (183, n =154) = 413.43, and the values from a CFA for non-music students χ^2 (183, n = 139) = 409.96 were aggregated to form baseline data χ^2 (366) = 823.39. The loadings were then constrained to be equal for each group and the analysis was run again. The χ^2 statistic increased by 110 to 933.69 and the degrees of freedom increased by 34 to 400. The large increase in chi-square value in relation to the increase in degrees of freedom indicated the items were biased in their assessment of student self-efficacy. It was therefore not tenable that the scale items were measuring music and non-music students equally and without bias. This finding was anticipated since non-music students were being asked to respond to music items.

To establish evidence of discriminant validity, the data were analyzed using multivariate analysis of variance (MANOVA). The results of this procedure indicated that scores generated by the MPSES were significantly different between musicians and non-musicians on self-efficacy as a composite variable, Λ = 0.63, F(4, 287) = 42.88, p <.0001. Univariate F-tests indicated significant differences between music and non-music students on each of the dependent variables (see Table 3). These results were supported by results from Tukey tests using a modified Bonferroni Type I error probability of α = .0125. Overall, there was a high level of power associated with the findings (1.00), a large effect size (f^2 = .59), and the corrected omega squared measure of association estimated 35% of the variance in the model was accounted for by the dependent variables.

Correlations were used to establish further evidence of validity. In a multi-method design, scores from the MPSES were correlated with the teacher ratings of students with very high and very low self-efficacy. MPSES scores were correlated with teacher ratings (r = .44) indicating a positive, yet moderate relationship. To determine whether the MPSES was investigating the same or similar types of self-efficacy as other scales, the correlations among scores on the MPSES (music), Bandura's CPSES (academic), and Pajares and Valiante's WSES (writing) self-efficacy scales were compared. The correlation between music and academic scores was r = .34 and the correlation between music and writing scores was r = .30. This comparison indicated the MPSES was measuring a different construct than the CPSES and the WSES. In contrast, the correlation between music and academic scores was r = .34 and the correlation between writing and academic scores was r = .67. This comparison indicated the difference between music self-efficacy and writing and academic self-efficacy establishing additional evidence of the MPSES's ability to generate scores that represent a unique form of self-efficacy.

Table 3
Univariate F-test Results of Sources of Self-Efficacy for Music and Non-Music Students

Source	*df*	*SS*	*MS*	*F*	*p*
Mastery Experience					
Music	1	4273774	4273774	117.51	< .0001
Error	290	10547191	36370		
Total	291	14820965			
Vicarious Experience					
Music	1	1672838	1672838	75.83	< .0001
Error	290	6397198	22059		
Total	291	8070036			
Verbal/Social Persuasion					
Music	1	4516404	4516404	149.91	< .0001
Error	290	8737118	30128		
Total	291	13253522			
Physiological State					
Music	1	2999725	2999725	163	< .0001
Error	290	5348480	18443		
Total	291	8348206			

Discussion

The primary purpose of this study was to develop and validate the Music Performance Self-Efficacy Scale. The development of the scale followed Bandura's (2006) guidelines with the exception of the response format. In this case, Bandura's 11-point Likert-type scale was replaced by a 0-100 continuous scale. I intentionally made this decision to position the MPSES alongside self-efficacy scales in other subject areas, to maximize the psychometric qualities of the scale, and to provide participants with a minimal amount of restriction in formulating their responses. It is important to note that although the internal consistency of responses from each section and the overall scale were very high, there was sufficient range in the inter-item correlations, item-total correlations, and correlations between the sources and the composite construct to suggest the items were not one dimensional. The items reflected a range of generality, strength, and level.

This study provided evidence supporting the MPSES's ability to measure self-efficacy in music performance. MANOVA and univariate *F*-test results indicated that the MPSES was able to distinguish between music and non-music students using scores from the total scale and from each source of self-efficacy section. This finding indicated the scale captured self-perceptions held by the music students that were not held by the non-music students. From a different perspective, comparisons of the correlation among the MPSES (music), the CPSES (academic), and the WSES (writing) indicated that the MPSES was measuring a related, but unique form of self-efficacy. The relationship between music and academic self-efficacy, and the relationship between music and writing self-efficacy were modest, while the relationship between academic and writing self-efficacy was moderately strong. Finally, a comparison of scores collected using different methods provided convergent evidence of the MPSES's ability to measure self-efficacy. The correlation between ratings supplied by the music teachers and scores on the MPSES was moderate (r = .44). This finding was encouraging given the subjective nature of the observations.

Confirmatory factor analysis provided evidence that the scale was generating data that fit Bandura's proposed model. The initial analysis provided results that were very close to accepted benchmarks. The modifications of removing items 9 and 13 not only improved the fit of the model, but also made it a shorter and more parsimonious assessment. This may be the most important finding from this study. This finding provides evidence that Bandura's model can be applied to music performance. Consequently, other models and frameworks based on

Bandura's framework may also be applied to music. Rather than *reinventing the wheel*, researchers interested in this topic might look to other studies for ideas and paradigms to explore self-efficacy.

Conclusion

The Music Performance Self-Efficacy Scale was constructed to measure the sources of self-efficacy in music performance. Although some self-efficacy measurements have been developed for specific events or contexts, the items in this scale were intentionally designed to be broad, allowing them to be applicable to different types of performing ensembles, different grade levels, different levels of music experience, and different times of the school year. This perspective follows Bandura's belief that the level of specificity in measuring self-efficacy should be consistent with the level of specificity to which one wishes to generalize. This broad and general scale of music performance self-efficacy was designed to reach broad and general conclusions. The unique feature of this scale is its ability to measure the sources of information that contribute to the development of self-efficacy beliefs.

This scale is a diagnostic tool. The target population for this scale is middle and high school music students. The results can be used to drive instructional choices based on the students' strengths and weaknesses. Educators may also use the data to evaluate the effectiveness of their own instructional practices and procedures. This scale can be used in a pretest-posttest design, or as a one-time "snapshot." To control consequential validity, it is essential that the results of this scale are kept confidential by the teacher and that they should not be used as a measure of achievement or academic grading.

It is my hope that the development and validation of this scale will provide researchers with a tool to pursue further investigations of self-efficacy in music performance. This study was conducted with one group of middle school students. Recommendations for future research would include carrying out a similar study with a larger and more diverse group of participants. It would also be beneficial to correlate the scores from this scale with music performance scores. This may be a particularly difficult task since music performance scores are often subjective and calculated in many different ways. Further examination of the differences in self-efficacy among students in various ensembles such as band, chorus, and string orchestra is another topic that warrants investment of time and energy. A final recommendation would be to establish scores for the sections, or sources of self-efficacy, that teachers can use as benchmarks. Teachers can then use these scores to compare their students' subtotal scores and determine the strengths and weaknesses of their students' self-efficacy in music performance. These efforts will result in a greater understanding of self-efficacy in music performance and improve student achievement by providing a balance of instruction in musical skills and self-perceptions in the classroom.

REFERENCES

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

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.

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.

Cattell, R. B. (1966). The meaning and strategic use of factor analysis. In R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology* (pp. 174-243). Chicago: Rand McNally.

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.

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.

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

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

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, for ProQuest Dissertations and 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.

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.

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. & McCormick, J. (2000). The contribution of motivational factors to instrumental performance in a music examination. *Research Studies in Music Education, 15*(1), 31-39.

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

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

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.

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.

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

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.

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

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

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.* (Unpublished doctoral dissertation) University of Iowa, Iowa City, Iowa.

Appendix A

Music Performance Self-Efficacy Scale

Identification Code:_____

Sources of 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 Strongly Disagree 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?_____

CHINESE ABSTRACT
中文摘要

音樂表演自我效能量表的設計與驗證

Michael S. Zelenak
Dixie M. Hollins High School—St. Petersburg, U.S.A

本研究的目的是完善音樂表演自我效能量表(MPSES)並驗證其有效性。本量表試圖測量中學生的音樂表演自我效能，從而驗證 Bandura 提出的四種自我效能资源，即：掌握經驗、替代經驗、言語說服、生理和情緒狀態。美國東南部的 293 個中學生參加了研究。研究結果顯示，通過本量表獲得的數據具有內在統一性($\alpha = .97$)。確認因素分析結果顯示了本研究獲得資料與 Bandura 模式之間的適合度 χ^2 (224) = 568.49，$p < .001$，*CFI* = .95，*RMSEA* = .07 及 *SRMR* = .04。這個結果可能是本研究最重要的發現。它證實了 Bandura 的四種自我效能資源能夠像在其他學科領域中一樣培育發展學生音樂的自我效能。其他能夠證實本量表能够作為測量音樂表演自我效能工具的證據還有：

1，音樂与非音樂學生的自我效能分數有很大的不同，$\Lambda = 0.63$，$F(4,287) = 42.88$，$p < .001$；
2，教師對學生的自我效能評估與量表資料有正相關性($r = .44$)；
3，本量表與其他量表所獲資料的相關對比所產生的判別式證據。

這些研究結果表明：本音樂表演自我效能量表是測量中學生音樂自我效能的有效與可靠的工具，研究者推薦將來有更大與更多樣的樣本來證實其有效性。