Music Achievement and Academic Achievement: Isolating the School as a Unit of Study

Marta Frey-Clark
Texas Tech University

Background

Music participation and academic achievement have long been of interest to educators, researchers and policy makers. The literature is replete with studies linking music participation to higher state assessment scores, grade point averages, and Standardized Achievement Test (SAT) scores. This growing body of research is not without its critics. In 2000, the *Journal of Aesthetic Education* published a special issue dedicated to the topic of arts and achievement. Several researchers in this issue focused on limitations of the extant research, ultimately cautioning against conclusions that music participation causes greater academic achievement (Vaughn & Winner, 2000; Winner & Hetland, 2000). However, in a climate of increasingly centralized curriculum and high-stakes testing, interest in links between music participation and academic achievement persists as the music community seeks to protect music's place in school curriculum.

The majority of researchers tested a correlational rather than causal relationship between music participation and academic achievement. Kinney and Forsythe (2005), for example, found that students who participated in arts-rich curriculum scored consistently higher on math, science, and citizenship achievement tests than students who participated in conventional curriculum. Thornton (2013) found that Pennsylvania music participants scored significantly higher in math and reading at every grade level of state testing (fifth, eighth, and eleventh grade), when compared with nonparticipants. Miksza (2007) compared achievement test scores of 5,335 students and found those who consistently participated in music from eighth through twelfth grade scored significantly higher than students with no history of music participation.

Researchers have consistently found a significant relationship between music participation and academic achievement; the nature of this relationship, however, remains unclear. Some researchers chose to test the influence of additional variables, such as socioeconomic status

(Catterall, Chapeau, & Iwanaga, 1999; Fitzpatrick, 2006; Miksza, 2007); others isolated particular characteristics of music participation—its length (Fitzpatrick, 2006; Miksza, 2007; Vaughn & Winner, 2000), quality (Johnson & Memmott, 2006), and type, as in general music, band, choir, or orchestra (Catterall et al., 1999; Elpus, 2013; Johnson & Memmott, 2006).

Socioeconomic status (SES) as an additional variable has received considerable attention in the literature on music and achievement. Researchers agree that SES is a powerful predictor of general academic achievement (Fowler & Walbert, 1991). As such, several researchers testing the relationship between musical involvement and academic achievement chose to include SES as an additional variable (Catterall et al., 1999; Fitzpatrick, 2006; Miksza, 2007). Their findings, however, are inconclusive. Some concluded the gap between high- and low-SES students remained constant over time, regardless of music participation (Miksza, 2007); others stated that, for instrumentalists, the gap shrank over time (Catterall et al., 1999); still others found that over the course of five years, test scores of low SES instrumentalists surpassed high SES non-instrumentalists (Fitzpatrick, 2006).

Researchers focused on instrumentalists as a distinct subset of the music population in response to a growing body of research concluding instrumentalists score higher than choir students on standardized tests (Elpus, 2013; Fitzpatrick, 2006; Kinney, 2008). In a study of 15,431 Ohio school students, Fitzpatrick (2006) found instrumentalists outperformed their non-instrumental counterparts across every grade level and subject on the Ohio Proficiency Test. Kinney (2008) examined band, choir, and nonmusic students and also found significantly higher scores for band students on the Ohio Proficiency Test. Catterall et al. (1999) analyzed data from the National Educational Longitudinal Survey, which consisted of ten years of math proficiency scores for over 25,000 secondary students. They found the number of twelfth-grade instrumentalists who met proficiency standards was significantly higher than non-instrumentalists. Elpus (2013) suggested that the achievement scores of instrumentalists alone perhaps drove the data throughout the body of research evidencing an achievement gap between music participants and non-participants. It is noteworthy, however, that researchers also found

would-be instrumentalists scored higher on achievement tests before receiving instrumental instruction (Elpus, 2013; Fitzpatrick, 2006; Kinney, 2008). They concluded that instrumental programs might attract higher achievers.

Socioeconomic status (SES) and type of music participation are relevant—though not exhaustive—variables in the study of music and academic achievement. A failure to account for all potentially relevant variables exposes this body of non-experimental research to criticism. Critics have called for rigorously designed experiments to investigate a causal relationship (Winner & Hetland, 2000). There remains, however, a dearth of experimental research on music participation and academic achievement. Furthermore, those few researchers who have conducted experimental studies found only modest or nonexistent relationships. Schellenberg (2004) tested four groups of six-year-olds. For one year, the researcher oversaw administration of Kodály lessons to one group, piano lessons to a second, theater lessons to a third, and no lessons to a fourth control group. Schellenberg administered an IQ test before and after the treatment. Kodály and piano students showed modest but significant gains in IQ. Costa-Giomi (2004) tested two groups of low-income fourth-grade students with no formal music education. Each student in the test group received an acoustic piano and three years of weekly piano lessons. Students in the control group received neither a piano nor lessons. Piano instruction had a positive effect on students' self-esteem and grades in music class, but did not impact standardized test scores.

Elpus (2013) acknowledged the limitations of correlational research and the shortage of experimental research and responded with a quasi-experimental study, addressing the issue of causality through regression analysis and the issue of omitted-variable bias through inclusion of a far greater number of variables than previous researchers. The data set consisted of college entrance exam scores for 15,630 students. Elpus (2013) compared music participants and nonparticipants and found the expected gap in academic achievement. When demographic variables were added to the model—gender, race/ethnicity, SES, and native language—the gap shrank. With the addition of prior academic achievement, Individualized Education Plan status,

time use, and attitudes towards school, the achievement gap virtually disappeared. Elpus (2013) suggested that variables affecting the choice to participate in music (specifically, instrumental music) also affect academic achievement. He concluded that selection bias had skewed previous studies, evidencing a link between music participation and academic achievement.

Elpus (2013) controlled for a large number of variables but did not allow for between-school comparisons. In the interest of reducing potential error, he controlled for between-school variation, acknowledging the elimination of the possibility of examining the school as a unit of study. Vaughn and Winner (2000) commented on the shortage of research on school-level variables, noting that schools are a worthy unit of study. They posited that the link between music participation and SAT scores could be a result of schools excelling both academically and musically, thus attracting excellent students to the music program. Johnson and Memmott (2006) also acknowledged the shortcoming of their study in failing to allow for between-school comparisons. Deviating from previous research linking music participation with academic achievement, Johnson and Memmott (2006) tested the potential of a link between music *achievemen*t and academic achievement. They compared standardized test scores of students from exemplary music programs with standardized test scores of students from deficient music programs. They found that students from exemplary music programs scored significantly higher on standardized tests. However, because they did not include the standardized test scores of the student body beyond the music program, they still did not isolate the school as a unit of study.

Educational researchers have reported that school-level variables such as cohesion (Stewart, 2008), school size, district size, and pupil-teacher ratio (Fowler & Walberg, 1991) are significantly related to academic achievement. Thus far, researchers examining achievement and *music* have focused on individual-level variables, such as student SES and type of music participation, to the exclusion of school-level variables.

If students from quality music programs academically outperform students from deficient music programs, and if school-level variables impact academic achievement, the question

remains: is there a link between a school's overall academic achievement and its overall music achievement? If so, is this relationship consistent for each type of music participation?

Research Design

For the purposes of this study, I operationalized music achievement as scores earned by middle school ensembles at the Concert & Sight-Reading Contest (C&SR) hosted by the University Interscholastic League (UIL) of Texas. The UIL is a statewide interscholastic organization created to provide educational competition in academics, athletics, and music. Each year, the UIL hosts a C&SR Contest in which secondary bands, choirs, and orchestras compete within prescribed regions. Per the C&SR Constitution and Contest Rules (University Interscholastic League, 2014), competing ensembles perform three prepared musical selections, comprising the concert portion of the contest, and one unfamiliar selection, comprising the sight-reading portion. Two panels of three judges adjudicate each portion. Judges assign ratings on a scale of 1 to 5, with 1 being the highest. Judges' scores are averaged to yield one contest score and one sight-reading score.

I operationalized academic achievement as passing rates on the State of Texas Assessment of Academic Readiness (STAAR), specifically, passing rates for seventh-grade and eighth-grade math and reading. Teachers administer STAAR tests for different subjects at different grade levels, but math and reading are consistently tested. The Texas Education Agency holds schools accountable for their passing rates on these two subjects alone (Adequate Yearly Progress, 2012), making them a key measure of a school's academic achievement.

I excluded sixth-grade passing rates because several schools participating in the UIL were junior high schools, comprised solely of seventh- and eighth-grade. Additionally, the UIL Constitution and Contest Rules require that ensembles be "comprised of a majority of 7th grade students and above" (2014), resulting in the isolation of seventh- and eighth-grade participants being the surest method of maintaining consistent measures.

Method

Middle schools from Regions 4, 8, and 21 of the UIL (N = 122) participated. These regions

constitute a moderate sample size and encompass a range of state assessment passing rates. I eliminated four schools from the data set: two because their STAAR scores were not available online, and two because they combined to form a single ensemble for the C&SR contest.

Raw data consisted of 2013 STAAR passing rates, retrieved from the website greatschools.org (2014); 2013 was the most recent year for which data were available. I summed passing rates for seventh-grade reading, seventh-grade math, eighth-grade reading, and eighth-grade math for each school to produce an aggregate STAAR passing rate. I then grouped schools according to aggregate STAAR passing rates. Group 1 (n = 44) consisted of high-performing schools, classified as those whose passing rates were more than half a standard deviation above the mean (M = 312.86, SD = 37.21). Group 2 (n = 46) consisted of average-performing schools whose passing rates fell within or equal to half a standard deviation of the mean. Group 3 (n = 32) consisted of low-performing schools whose passing rates were less than half of a standard deviation from the mean.

Raw data further consisted of C&SR scores for all middle school bands and choirs from schools in Regions 4, 8, and 21 of the UIL, and were retrieved from the website Texas UILforms.com (n.d.). For each ensemble, the website reports an average concert score and average sight-reading score in addition to individual scores assigned by each of the six judges. I bypassed average scores in favor of individual judges' scores, deeming discrepancies in individual ratings to be valuable data. I summed the six scores to produce a raw score for each ensemble. Because scores range from 1 to 5, 1 being the highest, raw scores then ranged from 6 to 30, 6 being the highest.

It is noteworthy that schools sent varied numbers and types of ensembles to the contest. There is no limit to the number of entries per school. The number of ensembles sent by a single school ranged from one to eight. A total of 174 bands—some from the same school—entered the contest, followed by 104 choirs and 28 orchestras. Thirteen of the schools that sent orchestras were academically low-performing schools. In practice, the number and variety of ensembles sent per school commonly reflects the size of the school's music program. I intended to examine

the overall quality of a school's music program, as opposed to its size or ensemble make-up, so I averaged raw scores for all ensembles from the same school to produce a single raw score for each school. I eliminated orchestras from the data set, deeming their numbers too small and unevenly distributed to yield significant results.

Results

Given the ordinal nature of the data, I conducted a Kruskal-Wallis one-way analysis of variance and found significant differences in C&SR scores of Group 1 (high-performing schools), Group 2 (average schools), and Group 3 (low-performing schools), $H(2, 122) = 10.18$, $p = .006$. A series of Mann-Whitney U tests identified that the significance lay between Groups 1 and 2, $Z(1, 90) = 2.47$, $p = .01$, and Groups 1 and 3, $Z(1, 76) = -2.99$, $p = .003$). No significance difference was found between Groups 2 and 3. Figure 1 shows the average C&SR scores for schools in each group.

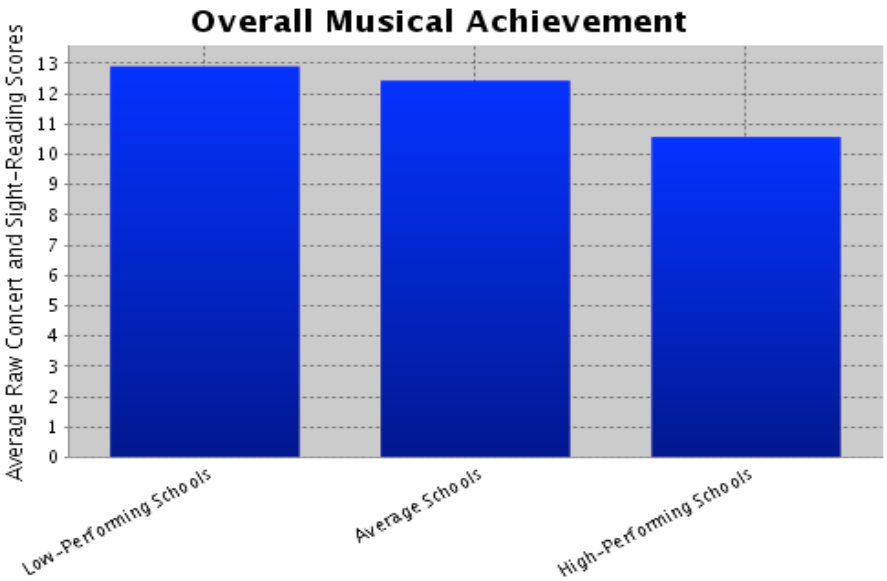

Figure 1. Average C&SR scores for all music ensembles by school STAAR passing rates. Note that a lower C&SR score indicates a greater musical achievement.

Next, I calculated Kruskal-Wallis analyses for bands and choirs separately, using average raw scores for each ensemble type. No significant difference was found between choir C&SR scores for high-performing, average, and low-performing schools.

I found a significant difference in C&SR band scores $H(2, 116) = 10.16$, $p = .006$. A series a Mann-Whitney U Tests revealed the significance lay between Groups 1 and 3, $Z(1, 72) = -3.18$, $p = .002$. No significant difference was found between Groups 1 and 2 or Groups 2 and 3. Figure 2 shows the average C&SR scores for bands and choirs.

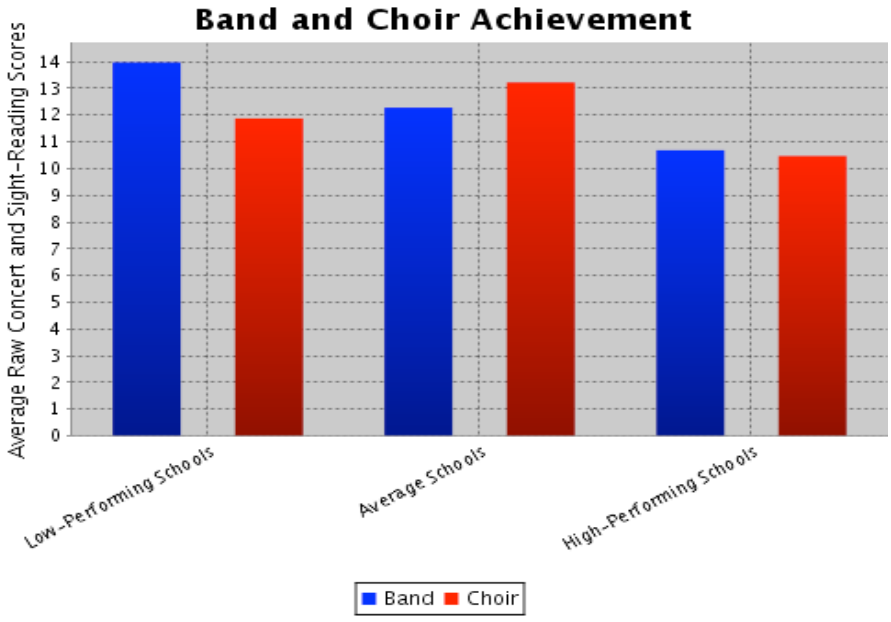

Figure 2. Average C&SR scores for bands and choirs by school STAAR passing rates. Note that a lower C&SR score indicates a greater musical achievement.

Discussion

It would appear that academically high-performing schools earned significantly higher Concert & Sight-Reading scores. This relationship could be the result of several variables omitted from this study. For example, if the demographic population of the music program closely resembles the population of the school, the relationship could be explained by individual-level variables. SES, for example, has been strongly linked to both academic achievement and music achievement (Catterall et al., 1999; Fitzpatrick, 2006; Miksza, 2007). If the school population and music population are similar, student SES could be a key variable

impacting this relationship.

This relationship may also be influenced by omitted school-level variables. The amount of money a district spends per child may impact both academic and musical achievement. School hiring practices could also play a role; schools that offer competitive salaries and hold teachers to high standards may excel in both arenas. Class size, school size, and district size may also have an impact.

School cohesion is another variable worthy of closer examination. Stewart (2008) described school cohesion as "the extent to which there is trust, shared expectations, and positive interactions among students, teachers, and administrators" (p. 190). Stewart (2008) found cohesion positively impacts student achievement. It is possible that school music programs have a two-way relationship with school cohesion. Additional research on music participation and school cohesion is warranted.

It is noteworthy that the relationship between academic and musical achievement was not consistent for each ensemble type. The difference between academically high- and low-performing schools virtually disappeared when choirs were isolated as a unit of study. Conversely, the difference between bands from high-performing and low-performing schools was significant at an alpha level of .01.

While several explanations could account for the difference between band and choir programs, one seems most likely: a disparity in resources available to high- and low-performing schools. Due to the very nature of instrumental study, band programs are more dependent on financial resources than their choral counterparts. There is cost involved in purchasing or renting an instrument, as well as maintaining it. Voices, on the other hand, are free. Private lessons are an added expense. One could argue that, in the absence of private lessons, vocal proficiency may be more attainable than instrumental proficiency—at least, at the middle school level. It is possible high-performing and low-performing schools represent a difference in resources that impacts band programs more than choir programs.

This study provides no basis for causal claims; it cannot support conclusions that schools

with strong academics produce strong musicians or that schools with strong music programs produce strong testers. The relationship could be causal, but an experimental or quasi-experimental study is necessary to test that possibility. The aforementioned omitted variables are equally likely to account for the link found between middle school STAAR scores and C&SR scores.

Readers are cautioned to avoid overgeneralizing the findings of this study. Participants were limited to a sample of three UIL regions. A replication including all 28 regions of the UIL would allow for statewide generalization. Replication that extends to other regions of the nation might allow for insights into the very nature of academic and musical achievement. An extension to high school C&SR scores and academic assessments would also augment the current findings. Finally, an expansion of the sample population might alter the definitions of high-performing, average, and low-performing schools, as these divisions were set according the mean and standard deviation of the data set. A shift in that definition may yield a subsequent shift in results.

Nonetheless, there was a significant difference in C&SR scores of high- and low-performing schools at an alpha level of .01. At the very least, it can be concluded that musical excellence did not detract from the academic performance of students in these 122 schools in East Texas. Local administrators and policy makers may benefit from this knowledge as they allocate limited time and resources for the education of Texas's youth.

Keywords

Music, musical achievement, academic achievement, band, choir, ensembles

Address for correspondence

Marta Frey-Clark, School of Music, Texas Tech University, Lubbock, Texas, 79409, marta.frey50@gmail.com

References

Catterall, J., Chapleau, R., & Iwanaga, J. (1999). Involvement in the arts and human development. In E. Fiske (Ed.), *Champions of change: The impact of the arts on learning*

(pp. 1-18). Washington, DC: The Arts Education Partnership and the President's Committee on the Arts and the Humanities.

Costa-Giomi, E. (2004). Effects of three years of piano instruction on children's academic achievement, school performance and self-esteem. *Psychology of Music*, *32*, 39-52. doi:10.1177/0305735604041491

Elpus, K. (2013). Is it the music or is it selection bias? A nationwide analysis of music and non-music students' SAT scores. *Journal of Research in Music Education, 61*, 175-194. doi:10.1177/0022429413485601

Fitzpatrick, K. R. (2006). The effect of instrumental music participation and socioeconomic status on Ohio fourth-, sixth-, and ninth-grade proficiency test performance. *Journal of Research in Music Education, 54*, 73-84. doi:10.1177/002242940605400106

Fowler, W. J., & Walberg, H. J. (1991). School size, characteristics, and outcomes. *Educational Evaluation and Policy Analysis, 13*, 189-202. doi:10.3102/01623737013002189

GreatSchools.org (2014). Find a school. Retrieved from http://www.greatschools.org/find-schools/.

Johnson, C. M., & Memmott, J. E. (2006). Examination of relationships between participation in school music programs of differing quality and standardized test results. *Journal of Research in Music Education, 54*, 293-307. doi:10.1177/002242940605400403

Kinney, D. W. (2008). Selected demographic variables, school music participation, and achievement test scores of urban middle school students. *Journal of Research in Music Education, 56*, 145-161. doi:10.1177/0022429408322530

Kinney, D. W., & Forsythe, J. L. (2005). The effects of the arts IMPACT curriculum upon student performance on the Ohio fourth-grade proficiency test. *Bulletin of the Council for Research in Music Education, 164*, 35-48.

Miksza, P. (2007). Music participation and socioeconomic status as correlates of change: A longitudinal analysis of academic achievement. *Bulletin of the Council for Research in Music Education, 172*, 41-58.

Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychology Science: A journal of the American Psychological Society, 15*, 511-514. doi:10.1111/j.0956-7976.2004.00711.x

Stewart, E. B. (2008). School structural characteristics, students effort, peer associations, and parental involvement: The influence of school- and individual-level factors on academic achievement. *Education and Urban Society, 40*(2), 179-204. doi:10.1177/0013124507304167

Texas Education Agency (2012). 2012 Adequate Yearly Progress (AYP): AYP Guide. Retrieved from http://ritter.tea.state.tx.us/ayp/2012/

Thornton, L. (2013). A comparison of state assessment scores between music and nonmusic students. *Update: Applications of Research in Music Education, 32*(1), 5-11. doi:10.1177/8755123313502339

University Interscholastic League (2014). Constitution and contest rules. Retrieved from http://www.uiltexas.org/policy/constitution.

Texas UILforms.com (n.d.) University Interscholastic League: Official concert & sightreading contest results. Retrieved from http://www.uilforms.com/bolt/csrrptUILpublic.asp

Vaughn, K., & Winner, E. (2000). SAT scores of students who study the arts: what we can and cannot conclude about the association. *Journal of Aesthetic Education, 34*(3/4), 77-89. doi:10.2307/3333638

Winner, E., & Hetland, L. (2000). The arts in education: Evaluating the evidence for a causal link. *Journal of Aesthetic Education, 34*(3/4), 3-10.