American Educational Research Journal
June 2017, Vol. 54, No. 3, pp. 399–443
DOI: 10.3102/0002831217701830

Does Participation in Music and Performing Arts Influence Child Development?

E. Michael Foster
University of Alabama at Birmingham
Jade V. Marcus Jenkins
University of California, Irvine

This article reconsiders the association between childhood arts participation and cognitive and developmental outcomes. Using data from a large, nationally representative sample with extensive covariates, we employ propensity score weighting to adjust comparisons of children who do and do not participate in arts education (music and performing arts lessons) to address potential confounding from selection into arts education. We examine a broad range of outcomes in adolescence and early adulthood (e.g., GPA, self-esteem, college attendance). Our results show that selection into arts education is at least as strong as any direct effect on outcomes, providing no support for the causal associations between arts participation and cognitive outcomes. We do find that arts education increases arts engagement during young adulthood.

KEYWORDS: achievement, arts education, child development, econometric analysis, longitudinal studies, motivation, secondary data analysis, self-concept

Introduction

The literature on arts learning and children's intellectual development in both childhood and adolescence shows consistent positive associations (Hallam, 2010). Of the arts, music education and participation is the most extensively studied and has the strongest associations (Southgate &

E. MICHAEL FOSTER (deceased), was a professor at the Department of Health Care Organization and Policy, University of Alabama at Birmingham. His work focused on policy issues related to children's development and on economic analysis of interventions and services for children with emotional and behavioral problems. This article is dedicated to his surviving wife and four children.

JADE V. MARCUS JENKINS, assistant professor, School of Education, University of California, Irvine, 2052 Education, Irvine, CA 92697; e-mail: *jvjenkin@uci.edu*. Her work focuses on early childhood development and public policy.

Roscigno, 2009). Music is also the most universally offered form of arts education, available in more than 90% of public schools nationwide (Parsad & Spiegelman, 2012). Creating music, especially at the professional level, is a highly complex task that engages multiple brain systems (Munte, Nager, Beiss, Schroeder, & Erne, 2003). Some of these systems are also involved in aspects of learning, such as reading (Anvari, Trainor, Woodside, & Levy, 2002; Gardiner, Fox, Knowles, & Jeffrey, 1996; Magne, Schon, & Besson, 2006).

An association between music and other arts education and cognitive development, therefore, is plausible. However, a key issue for public policy is whether this relationship is causal—that is, whether arts education actually improves learning and brain development. Efforts to expose more children to music, for example, at an early age presume such a relationship. Perhaps best known is the "Mozart for Babies" program. In 1998, then-Governor Zell Miller of Georgia allocated $105,000 to provide newborns with a classical music CD. The presumption was that listening to the music would permanently improve the child's cognitive development. This claim, however, was based on (only) an association between music exposure and child development.

But is the association causal? Will increasing exposure to music or other types of arts education improve key child development outcomes like math or reading achievement? The answer at this point is largely unknown. Existing research is at best speculative. It is accepted that music rests on underlying mathematical principles (Vaughn, 2000), but whether exposure to music would improve children's understanding of these principles and application to more explicitly mathematical tasks is presumed but not demonstrated. The link between music education and spatial reasoning is among the strongest associations that researchers have documented (Hallam, 2010). In addition, experimental (e.g., Greenfader, Brouillette, & Farkas, 2015; Moreno et al., 2009) and associational (e.g., Piro & Ortiz, 2009) studies also show a positive relationship between music education and language and reading skills. Yet meta-analyses of experimental studies of music education interventions highlight substantial limitations in that research, such as a failure to generalize findings between studies, and small sample sizes (Hetland, 2000; Southgate & Roscigno, 2009; Winner & Cooper, 2000). Indeed, the body of empirical studies on the effects of arts education on child development outcomes such as language, reading, and math skills show a mix of positive and null findings (see Table 1 for more detail).

Moving from association to causal claims is always challenging. Methodological research indicates that doing so—outside of random assignment to treatment (e.g., music lessons)—always involves an assumption that cannot be formally tested. One such assumption of nonexperimental studies is "exchangeability"—that the outcomes that children who do not participate in music lessons (i.e., untreated) would experience were they to take music

Table 1

Review of Recent Studies Examining the Effects of Arts Participation on Child and Adolescent Outcomes by Overall Findings

Positive

Year	Authors	Title	Exposure	Data/Country	Method	Outcome	Covariates	Findings
2016	Winsler, A., Gara, T., Alegrado, A., Castro, S., Tavassolie, T.	Selection into, and academic benefits from, arts-related elective courses in middle-school among low-income, ethnically diverse youth	Participation in arts courses (music, drama, visual art, dance) during middle school	31,322 Miami public school children, followed from pre-k entry through 8th grade; USA	OLS and logistic regression	GPA, reading and math achievement, retention, suspension, absenteeism in middle school	Gender, race, free or reduced price lunch, English proficiency, prior achievement at grade 5 and age 4	Children with stronger skills at pre-k more likely to enroll in arts. Students who took arts courses had higher GPAs, math, and reading scores, lower odds of suspension
2011	Moreno, S., Bialystok, E., Barac, R., Schellenberg, E.G., Capeda, N.J., Chau. T.	Short-term music training enhances verbal intelligence and executive function	Random assignment to 20-day computerized music training program vs. computerized visual arts training program	71 four- to six-year-old children; Canada	Experimental design; ANOVA with observed groups	Verbal intelligence, executive function	None	Children who received music training showed improvement in verbal intelligence and executive function
2010	Helmrich, B.H.	Window of opportunity? Adolescence, music, and algebra	Formal instrumental or choral music instruction during middle school	6,026 adolescents enrolled in Maryland High Schools; USA	ANCOVA between observed groups	Eighth and 9th grade introductory algebra scores	Fifth grade math ability, race	Students in instrumental music courses had higher algebra scores than choral students and no-music instruction students
2009	Piro, J.M., Ortiz, C.	The effect of piano lessons on the vocabulary and verbal sequencing skills of primary grade students	Public school with music instruction vs. school without music instruction	101 second-grade students; USA	ANCOVA between observed groups	Vocabulary; verbal sequencing	Gender, race, income, household net worth, yearly expenditures, Title I school, licensed teachers, attendance rates, percentage of children receiving free lunch, reading score	Children exposed in music curriculum schools performed better on vocabulary and verbal sequencing tests

(continued)

Table 1 (continued)

Year	Authors	Title	Exposure	Data/Country	Method	Outcome	Covariates	Findings
2009	Moreno, S., Marques, C., Santos, A., Santos, M. Castro, S.L., Besson, M.	Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity	Random assignment to in-school musical or painting lessons 2x/ week for 6 months	32 eight-year-old children; Portugal	Experimental design; ANOVA with compared exposures. Analysis of brain wave patterns using EEG.	Reading ability, IQ, working memory, brain wave patterns and pitch discrimination	Sex, socioeconomic status, prior musical training	Music students had enhanced reading and pitch discrimination in speech
2006	Schellenberg, E. G.	Long-term positive associations between music lessons and IQ	Participation in group or private music lessons	Two samples: 147 six- to eleven-year-old children, 140 college undergraduates; Canada	OLS regression	Intelligence, academic ability, and social adjustment	Family income, linguistic background, country of birth, parents' education	Significant adjusted associations across broad range of outcomes
2004	Schellenberg, E. G.	Music lessons enhance IQ	Participation in music lessons (keyboard, voice), participation in drama lessons (36 weeks)	144 six-year-old children; Canada	Quasi-experimental design-non equivalent group design	IQ and academic achievement	Age, family income	IQ increases for music students
2003	Ho, Y., Cheung, M., Chan, A.S.	Music training improves verbal but not visual memory: Cross sectional and longitudinal explorations in children	Participation in instrumental music training at school vs. those who did not	90 male participants aged 6 to 15; Hong Kong	ANOVA between observed groups	Verbal and visual memory, intelligence	Age, education level, father and mother's education level, family income	Musical lessons group showed improved verbal learning and retention abilities
2002	Anvari, S.H., Trainor, L.J., Woodside, J., Levy, B.A.	Relations among musical skills, phonological processing, and early reading ability in preschool children	Children were administered a set of music tasks focused on rhythm, melody, and chord processing	100 four- to five-year-old children; Canada	OLS regression	Reading ability	Age, music performance score, phonemic awareness, vocabulary, math and memory score	Music performance score correlated with reading skills and phonological awareness
2001	Fisher, D.	Early language learning with and without music	Random assignment to classrooms where teachers used music in daily activities	80 bilingual children, kindergarten and first grade; USA	Experimental design; ANOVA with compared exposures	Reading achievement	Income	Students in music classroom had improved oral language and reading scores

(continued)

Table 1 **(continued)**

Year	Authors	Title	Exposure	Data/Country	Method	Outcome	Covariates	Findings	
1999	Catterall, J., Chapleau, R., Iwanaga, J.	Involvement in the arts and human development: General involvement and intensive involvement in music and theatre arts	School-based music activity. Supplemental analyses consider the duration of involvement.	National Educational Longitudinal Survey (1988 cohort) (n = 13,327); USA	Descriptive comparisons only	Standardized tests and academic achievement	Socioeconomic status (family education level, income, and type of job(s) held by parents)	Correlation between involvement in the arts and academic Success in high school, and between music and math achievement	Positive
1994	Morrison, S. J.	Music students and academic growth	School-based music activity	National Educational Longitudinal Survey (1988 cohort) (n = 13,327); USA	Descriptive comparisons only	Academic achievement	Socioeconomic status	Substantial correlation between SES and music involvement, and between music involvement and achievement	
1975	Hurwitz, I., Wolff, P.H., Bortnick, B.D., Kokas, K.	Nonmusical effects of the Kodály music curriculum in primary grade children	Random assignment to Kodály Music Training Program (Instruction for 40 min/day for 7 months)	40 middle-class children ages 5–7 years; USA	Experimental design with ANOVA	Sensorimotor and verbal perceptual sequencing, verbal intelligence	Matched on age, IQ and social class (Hollingshead index)	Improved performance in temporal and spatial tasks and reading skills	
2013	Mehr, S.A., Brady, A.M., Katz, R.C., Spelke, E.S.	Two randomized trials provide no consistent evidence for nonmusical cognitive benefits of brief preschool music enrichment	Random assignment to brief series of parent-child music classes vs. similar but non-musical form of arts instruction (Study 1), vs. a no-treatment control (Study 2)	91 preschool children and their parents (total both experiments); USA	Experimental design with ANOVA with compared exposures	Map use/navigation, 2-D visual form analysis, numerical quantity discrimination, and receptive vocabulary	Family income, parent's total work hours, music aptitude	Children provided with music classes performed no better than those with visual arts or no classes on any assessment.	Null
2013	Elpus, K.	Is it the music or is it selection bias? A nationwide analysis of music and non-music students' SAT scores	High school course credit in music	Education Longitudinal Study, 2002. (n = 15,630); USA	School fixed effects	SAT score, standardized math exam	Race, prior achievement, socioeconomic status, time use, attitudes towards school, family composition	Music students did not out-perform non-music students when adjusting analyses for school-level factors	

(continued)

Table 1 **(continued)**

Year	Authors	Title	Exposure	Data/Country	Method	Outcome	Covariates	Findings
2012	Rickard, N.S., Bambrick, C.J., Gill, A.	Absence of widespread psychosocial and cognitive effects of school-based music instruction in 10–13-year-old students	Random assignment to 6 month classroom-based music program vs. a drama, or visual arts program (Study 1); Music education program in psychosocial well-being classes vs. drama program (Study 2)	227 ten- to thirteen-year-old private school students; Australia	Experimental design; ANOVA with compared exposures	Verbal memory; intelligence; self-esteem; attitudes to school and motivation; and engagement with classes	Socioeconomic status	No significant differences between any program exposure contrasts
2011	Schellenberg, E. G.	Examining the association between music lessons and intelligence	Musically trained children vs. untrained children	106 nine- to twelve-year old children; Canada	ANOVA and MANOVA between observed groups	Several measures of intelligence and executive function	Parental education, involvement in out of school activities, native English speaker, income	Music training was not related to performance on executive function test
2000	Schneider, T. W., Klotz, J.	The impact of music education and athletic participation on academic achievement	Participation in band or choir, athletics, or neither activity	346 fifth- and sixth-grade students; USA	MANOVA between observed groups	Reading, language, and mathematics achievement tests	None	No significant differences between musicians and non musicians
2000	Winner, E. Cooper, M.	Mute those claims: No evidence (yet) for a causal link between arts study and academic achievement	Meta analysis of experimental and non-experimental studies	Varies	Meta-analysis	Several measures of verbal and math ability	None	No consistent evidence of benefits of music on child ability
1999	Costa-Giomi, E.	The effects of three years of piano instruction on children's cognitive development	Random assignment to piano instruction once a week for three years	117 fourth-grade students; Canada	Experimental design with ANOVA	Cognitive and spatial abilities	Gender, income, single or two-parent family, and parental employment	No difference between groups at the end of 3 year treatment; some improvements for treated group in year 2

(continued)

Table 1 **(continued)**

Year	Authors	Title	Exposure	Data/Country	Method	Outcome	Covariates	Findings	
1993	Zulauf, M.	Three-year experiment in extended music teaching in Switzerland: The different effects observed in a group of French speaking pupils	Random assignment to treatment classrooms who received 5 lessons per week for 3 years; Control 1-2 lessons per week (status quo for Swiss curriculum)	397 sixth-grade students; Switzerland	Experimental design	Academic achievement	Sex	No differences between treatment and control on all academic outcomes	Null
2015	Greenfader, C., Brouillette, L., Farkas, G.	Effects of a performing arts program on the oral language skills of young English learners	Random assignment to 28 50-minute sessions of drama and creative movement of English literacy lessons	5,1240 five- to eight-year-old children in low-income public schools; USA	Experimental design; OLS regression	Score on English-language proficiency exam	Gender, race, parent education, grade, prior achievement	Treated students with low English language speaking skills improved oral language skills, but not higher skilled students	Mixed
2010	Schellenberg, E. G., Moreno, S.	Music lessons, pitch processing, and g	Students who participated in music training vs. students who did not	40 undergraduate students; Canada	MANOVA between observed groups	Pitch processing and general intelligence	None	Music trained group had improved pitch processing for tones	
2009	Wetter, O. E., Koerner, F., Schwaninger, A.	Does musical training improve school performance?	Children practicing music at home (piano lessons) vs. those who do not	134 third- to sixth-grade children; Switzerland	ANOVA between observed groups	School performance	Child's grade, average income of parent's occupation, gender	Children involved in music had higher grades and better maintained those grades over time.	
2009	Southgate, D. E., Roscigno, V. J.	The impact of music on child and adolescent achievement	Music participation: in school, outside of school (lessons), and parental involvement (such as concert attendance)	Early Childhood Longitudinal Survey (Kindergarten Cohort); National Educational Longitudinal Survey (1988 cohort) (*n* = 4,376, 7,781); USA	OLS and logistic regression	Math and reading achievement	Race/ethnicity; family structure; number of siblings; gender and whether the home includes 50 or more books. Family socioeconomic measure constructed from income, parental education and parental occupation.	Music in school predictive of math achievement for children and not adolescents, but not music outside of school. Music in school predicted reading achievement, with larger effects for children than adolescents. Social class differences in outside school music participation	

(continued)

Table 1 **(continued)**

Year	Authors	Title	Exposure	Data/Country	Method	Outcome	Covariates	Findings
2008	Forgeard, M., Winner, E., Norton, A., Schlaug, G.	Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning	Children who received instrumental music training for three years vs. children who did not	59 eight- to eleven-year-old children; USA	MANCOVA between observed groups	Several measures of intelligence	Parental education	Music involvement related to fine motor ability and melody discrimination, verbal ability, but not spatial skills, overall intelligence, phonemic awareness
2006	Fitzpatrick, K.	The effect of instrumental music participation and socioeconomic status on Ohio fourth-, sixth-, and ninth-grade proficiency test performance	Students who participated in instrumental music in school	15,431 fourth-, sixth, and ninth-grade students; USA	ANOVA between observed groups	Citizenship, math, science, and reading test scores	Socioeconomic status	Instrumental students outperformed noninstrumental students in all subjects, but instrumental students outperformed noninstrumental students before instrumental lessons began
2005	Schlaug, G., Norton, A., Overy, K., Winner, E.	Effects of music training on the child's brain and cognitive development	Children who received instrumental music training versus those who did not	75 children 5 to 7 years old (50 musicians, 25 not); USA	ANOVA between observed groups	Structural and functional brain measures (MRI, FMRI), several measures of intelligence	Age, SES, verbal IQ	More grey matter volume for music training group and improved scores on some domains (vocabulary, audiation), but not others. Some differences in activation in neural systems
2001	Barber, B.L., Eccles, J.S., Stone, M.	What ever happened to the Jock, the Brain, and the Princess? Young adult pathways linked to adolescent activity involvement and social identity	Participation in school band, drama, or dance during high school	900 twelfth-grade students from Michigan Study of Life Transitions; USA	MANOVA between observed groups	Substance use, social adjustment, self-esteem, school involvement, educational attainment	Gender, math and verbal ability, mother's education	Performing arts participation predicted more years of education, increases in drinking between ages 18 and 21, higher rates of suicide attempts and psychologist visits by 24 years

Mixed

(continued)

Table 1 **(continued)**

Year	Authors	Title	Exposure	Data/Country	Method	Outcome	Covariates	Findings
2000	Bilhartz, T.D., Bruhn, R.A., Olson, J.E.	The effect of early music training on child cognitive development	Treatment group received music curriculum, 1x/week for 30 weeks that included parental involvement	71 four- to six-year-old children; USA	Experimental design; ANOVA between groups	Several measures of intelligence	Sex, ethnicity, parental education, economic class	Improved performance in spatial-temporal reasoning scores
2000	Vaughn, K.	Music and mathematics: Modest support for the oft-claimed relationship	Meta analysis of experimental and non-experimental studies	Varies	Meta-analysis	Several measures of mathematics achievement	None	Some positive, small associations between music and math achievement, mostly for children with special needs
1999	Eccles, J. S., Barber, B.L.	Student council, volunteering, basketball, or marching band: What kind of extracurricular involvement matters?	Participation in school band, drama, or dance during 10th grade	1,259 twelfth-grade students from Michigan Study of Life Transitions; USA	MANOVA between observed groups	GPA, school engagement, general ability, college attendance, substance use, social adjustment	Gender, mother's education	Participation in performing arts related to greater liking of school, higher 12th-grade GPA, and a greater likelihood of attending college at age 21. Reduced drinking behavior during high school for males only
1998	Chan, A.S., Ho, Y., and Cheung, M.	Music training improves verbal memory	Students who received 6 years of musical instrument training before age 12 vs. those who did not	60 female college students (30 trained, 30 not); Hong Kong	ANOVA between observed groups	Verbal and visual memory	Comparison group matched to treated group on age, GPA, & years of education	Musically trained adults had better verbal memory than non-musically trained. No difference on visual memory

lessons can be represented by the experiences of *otherwise comparable* children who *do* take music lessons (i.e., treated). In other words, the researcher assumes that the children being compared are similar in every way *except* their exposure to music lessons.

In most instances this type of research—without an experimental control group—compares children who were and were not involved in the arts and adjusts that comparison for a limited set of child and family characteristics. In this case, one assumes that assignment to treatment is ignorable; there exist no systematic relationships between selecting into arts lessons and other observed or *unobserved* child or family characteristics (exchangeability holds). Any residual difference in outcomes between the two groups is described as the "effect" of music education.[1]

The principal threat to causal inference in this kind of analysis is that children exposed to arts education differ from "otherwise comparable" children in ways *other than arts education*. In that case, the adjusted association between arts education and developmental outcomes reflects the effect of the former as well as these other factors, known as confounders. Exchangeability assumes the list of confounders is complete. One can assess this assumption in light of the list of variables included in the analysis and what one knows from other research about the predictors of music education participation. To use an example from another area of research, researchers have long been interested in the effect of breastfeeding on a child's intelligence. That research involves comparing children who were and were not breastfed. Most of that research does not include the mother's IQ in the statistical adjustment. As a result, comparisons of children who were and were not breastfed does not involve *otherwise comparable* children—such comparisons confuse the effect of breastfeeding with that of mothers' IQ.

How plausible is the claim of exchangeability or no unobserved confounding in the literature on the effects of different types of arts education? To some extent, all researchers recognize the potential for confounding. With respect to music education, Hallam (2010) notes that "one of the difficulties with this research, however, is that participating in musical activities may be related to other factors which promote academic attainment, for instance, having supportive parents and a home environment conducive to studying" (p. 277). A recent study notes that "without assessing social class variations, or controlling for other educational resources at children's disposal, conclusions regarding the impact of music participation are tentative at best" (Southgate & Roscigno, 2009, p. 7). Like other studies, however, comparisons of those who are and are not involved in arts education are very limited. Many authors do not account for socioeconomic status (SES) or, if they do, include a single measure of SES—one that combines family income, parental education and parental occupation. If among families with the same level of composite SES, differences in the constituent characteristics (e.g., income) remain, the effects of these factors on child outcomes could very well be misattributed to arts education.

This article recognizes the potential for confounding from both family SES and from the child's own characteristics. Such self-selection seems intuitive—children with more natural abilities may be more successful at music and spend more time engaged in music education. The association between music education and mathematical or verbal reasoning may be two-way: The latter may cause as well as reflect the former. The same may be true for other types of arts education; children who exhibit more prosocial behavior or have higher perceptions of self-efficacy may be more likely to engage in the performing arts. Our analyses test for these associations.

This study examines the effect of children's music or performing arts education (defined as lessons) on their development. This issue seems to be one where an association is not enough. A causal estimate has a variety of policy implications; broader cognitive benefits strengthen the case for public support of arts education; schools might substitute instructional time in academic subjects with arts. Indeed, we were funded by the National Endowment for the Arts (NEA) to examine this issue specifically. The original request for applications indicated that there was a lot of evidence of some association, and the NEA commissioned us to determine whether this relationship was causal. This endeavor need not indicate that the value of arts education should be determined by whether or not it influences child outcomes; participation in the arts is culturally valued in and of itself, and engaging in these activities may promote positive youth development regardless of its influence (or lack thereof) on any cognitive domains (Chappell & Cahnmann-Taylor, 2013; Larson, 2000; Lerner, Dowling, & Anderson, 2003). However, examining whether these associations are real is an important scientific question.

We reconsider the association between arts education and child development outcomes, improving on prior work in several ways. Our data come from a large, nationally representative dataset, the Panel Study of Income Dynamics Child Development Supplement. We developed our analysis plan to remove potential confounding—that is, to make the assumption of exchangeability as plausible as possible for these analyses. First, we selected a more extensive set of covariates than has been used in prior research. Second, we used improved methodology for adjusting comparisons of children who are and are not engaged in arts education for these covariates (propensity score weighting). In addition, we address missing data with multiple imputation methods to maximize the use of available data under plausible assumptions. We look at three different types of arts education exposures: lessons in musical arts, lessons in performing arts, and use of a musical instrument in the home. We examine the effects of these exposures on a broad range of outcomes in both adolescence and early adulthood including academic ability, self-esteem, high school completion, and future participation in the arts.

Background

Of the arts, music education is the most extensively studied. It is also the most universally offered form of arts education in public schools; national school survey data indicate that 94% of elementary schools and 91% of secondary schools offer opportunities for music education (Parsad & Spiegelman, 2012). As reviewed in Hallam (2010) and Schellenberg (2011), the literature documenting an association between music education and intelligence is substantial. Both reviews note that the associations between music training and cognitive abilities are quite general and extend across a wide variety of developmental domains. Because of this generality, it may not be surprising that the empirical research examining the relationship between arts education and key child development and achievement outcomes is equivocal.

Is There Evidence That Music Education Influences Children's Outcomes?

We summarize over 30 articles examining the relationship between music and development in Table 1. Studies are grouped by their primary finding. The table makes clear that the causal effects of arts education (though the reviewed studies look primarily at music education) on outcomes such as language, reading, and math skills are ambiguous. Approximately one-third of the studies find positive effects, another third find no effects, and the remaining third produce mixed results. For instance, Southgate and Roscigno (2009) find that music lessons in school were predictive of math achievement for children, but not for adolescents. Schellenberg and Moreno (2010) find a positive relationship between music training and pitch processing, but not intelligence. Pitch processing is clearly a more proximal effect of music training, but intelligence would require much broader and persistent transfer of skills between music and cognition. Similarly, Bilhartz et al.'s (1999) experiment finds an association between students who received a music curriculum and improved spatial-temporal reasoning, but no effects on other measures of intelligence used in the study. A review of meta-analytic studies came to a similar conclusion: Some causal evidence may exist for the influence of the arts on very specific cognitive tasks, but for many developmental outcomes there is no reliable causal support (Hetland & Winner, 2001).

A wide range of explanations exist for why music education should improve cognitive function including physical changes within the brain itself (for an excellent review, see Hallam, 2010). Because speech and music share several neural processing systems, musical experiences may improve phonological awareness (Anvari et al., 2002). For example, music can improve linguistic pitch processing and the sensory encoding of sound, which can affect language learning (Moreno et al., 2009; Patel, 2009; Wong, Skoe, Russo, Dees, & Kraus, 2007). In turn, this improves perception of language, which can then transfer directly to early reading skills and reading comprehension (Anvari et al., 2002; Gardiner et al., 1996; Magne et al., 2006).

Even if arts education did not impact academic performance per se, a large body of evidence in developmental psychology highlights the importance of involvement in the arts, and in constructed leisure activities more generally, in promoting positive youth development during adolescence and early adulthood (Eccles & Barber, 1999; Larson, 2000). Activities like musical and performing arts help to structure a child's peer group, facilitate social relatedness, form identity, express talents, and achieve positive recognition, and may help steer adolescents away from risky behaviors like skipping school and using drugs (Barber, Eccles, & Stone, 2001; Eccles, Barber, Stone, & Hunt, 2003; Fredricks et al., 2002; Fredricks & Eccles, 2006; Mahoney, Cairns, & Farmer, 2003). Engaging in an activity can facilitate one's intrinsic desire to learn, and this is associated with long-term overall well-being from continued engagement (Deci & Ryan, 2008; Larson, 2000; Lerner et al., 2003). Researchers in this area posit that when children have a "spark"—a passion for a self-identified interest or skill—this purpose and direction allows them to thrive and make positive contributions in community and civic life (Bobek, Zaff, Li, & Lerner, 2009; Lerner et al., 2005; Scales, Benson, & Roehlkepartain, 2011). For these reasons alone, it is likely in the interest of public schools to offer educational opportunities in the arts.

Causal inference requires a balance between an understanding of the theoretical mechanism linking exposures to their effects and an understanding of the causes of the exposure itself. Our focus here is on the latter. We use the literature in this area to identify child and family characteristics that are potential confounders—factors that influence both the selection into the arts as well as the key developmental outcomes to which arts education is putatively related. Understanding the causes of exposure to arts education is the primary concern underlying economic theory, and these concerns about unobserved variables and causal relationships shape all subsequent empirical work in the present study (Becker & Tomes, 1994).

Does the Existing Literature Provide Causal Estimates?

Recognizing potential confounding, existing studies generally employ regression to adjust for a set of observed factors. To interpret these adjusted associations as causal relationships, the list of covariates has to be complete, and there must be no confounding from *unobserved* variables. We could expect a range of variables to influence a family's choice to enroll a child in arts education: income, parent's ability and education, and child and community characteristics.

In our summary of the literature in Table 1, we assess whether prior research in this area adjusted for potential confounders. Looking at the covariates column, one can see that the analyses often include a meager set of controls, if any covariate adjustments are made at all. Confounding from unmeasured factors is clearly an issue with the prior research on arts education.

As noted, a prominent confounder suggested by economics involves the child's ability. Parents may make greater investments in children who are already more capable, or more capable children may be more motivated to participate in the arts. This possibility has also received attention from education and psychology researchers (Fitzpatrick, 2006; Kinney, 2010). Indeed, Schellenberg (2011) argues that "the vast bulk of the available literature can be explained simply: High-functioning children are more likely than other children to take music lessons, and to perform well on virtually any test they take" (p. 285). Yet reviewing the existing studies shows that children's ability measured prior to the exposure is rarely included in the analyses.

Many observable characteristics of children and families are associated with music participation. Indeed, nationally and regionally representative studies of high school students shows that children who participate in band, choir, and orchestra are primarily from higher SES backgrounds, are white and native English speakers, have the highest standardized test scores and GPAs, and their parents have advanced degrees (Eccles & Barber, 1999; Elpus & Abril, 2011; Feldman & Matjasko, 2007). Moreover, these factors, especially SES and ability, predict student's persistence and success in musical arts (Corenblum & Marshall, 1998; Kinney, 2010; Klinedinst, 1991), and are likely compounded by differential access to music education in public schools based on school-level SES (Parsad & Spiegelman, 2012; Winsler, Gara, Alegrado, Castro, & Tavassolie, 2016). These factors are inextricably linked to children's achievement and well-being. Recently, Elpus (2013) provided the first empirical evidence that the observed test score advantages for music students was not due to music education but to selection bias.

Additionally, the unobserved characteristics of parents could confound the observed relationship between a given investment and children's outcomes. For example, some parents may be "high-investing" regardless of how they perceive their child's ability. In this case, the relationship between exposure to investments, like music, and children's academic success would be confounded by the unobserved characteristics of their parents. In line with the research on music lessons mentioned above, these children would perform well at many tasks because their parents created a strong and perhaps diverse child investment portfolio. This makes it quite challenging to identify the *unique* causal effect of any one investment (i.e., arts education).

This area of research in general—this article being no exception— suffers from technical problems as well. As described in Appendix A in the online version of the journal, the standard tool applied in this research, regression, suffers from limitations. Regression makes a set of assumptions that have many benefits when examining a broad range of questions. However, if these assumptions are incorrect, the estimated effect of music education may be incorrect. For instance, regression assumes that all relationships are linear. As an example, an increase in SES from very low to low is presumed to have the same effect on outcomes as that from very high to higher still. Similarly, regression assumes the relationships between covariates and music participation are linear. To the extent this assumption is incorrect, the effect of music education is likely confounded: The regression model does not properly adjust comparisons for the covariates involved. Thus, the association between arts education and achievement can reflect not only confounding but also model misspecification. While these issues seem technical, they can easily create associations where no cause and effect relationship exists. The study methodology was designed to address these issues.

Methodology

Our study examines the effect of three different types of arts education exposures on a broad range of outcomes in both adolescence and early adulthood using a large, representative sample of children. The goal of our methodology is to remove potential confounding to obtain minimally biased estimates of the effect of arts education on development. We use an extensive set of covariates and address selection bias from family and child characteristics using propensity score weighting and address missing data with multiple imputation. Details are described below.

Data

The Panel Study of Income Dynamics (PSID) is a nationally representative household panel survey, collecting data since 1968. The original study contained a sample of 18,000 individuals living in 5,000 families in the United States. Data on these individuals and their descendants have been collected through successive waves and include extensive information related to health, occupation, income, education, and many other topics. This analysis relies on three components of the PSID: Main Family Interview, Child Development Supplement (CDS), and Transition Into Adulthood (TA). The PSID has maintained reasonable response rates over time. The representativeness of the CDS in particular has received extensive analyses, and the data included in the CDS and the TA studies represent the experiences of children ages 0 to 12 in 1997 (Hofferth, Davis-Kean, Davis, & Finkelstein, 1998). Figure 1 describes the data files used in the analysis and the relationship among them. Appendix B in the online version of the journal provides a comprehensive list of the variables we use, variable description, the PSID question number(s) from which they were created and question wording, and references for each instrument and assessment.

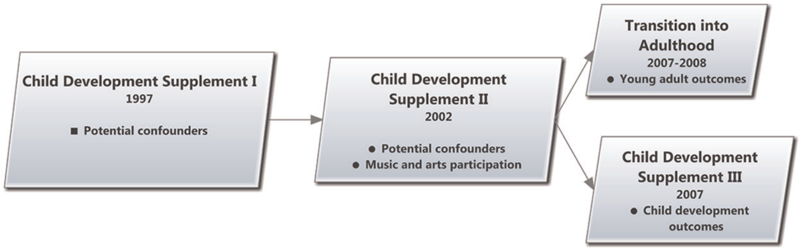

Figure 1. **Chronological order of data files used in analysis and the relationship among them.**

Main Family Interview

On a regular basis, one person per family is interviewed. Information is collected about each family member, but the most detail is collected on the heads of household. This component of the PSID has been conducted annually from 1968 through 1997, and biennially thereafter. Topics include employment, income, health, and education. We used the main interview data from 1997-2002 to capture information about the parents and family.

Child Development Supplement

The first CDS wave (CDS I) was implemented in 1997 when children were ages birth to 12 years, with follow-up interviews in 2002/2003 (CDS II) when children were ages 5 to 18, and 2007/2008 (CDS III; ages 10 to 19). Interviews were conducted with the child and the child's caregiver(s), absent parent, teacher, and school administrator. These data include information related to health and well-being, academic achievement, relationships, and time use. We used all three waves of the CDS data in the present study.

Transition Into Adulthood

This study interviews children from the CDS cohort who were between ages 18 and 24 at the start of the study in 2005, and biennially thereafter. This captures information about life changes and experiences during the transition to adulthood including time use, employment, income, education, and career goals. We used the 2007 TA data in this study.

Definition of Exposures

These analyses used children's arts exposure (or participation) at the second wave (CDS II) to predict child outcomes at the third wave (CDS

III) and during the transition to adulthood (TA 2007), using covariates measured in the first and second waves (CDS I and II) and from the corresponding main interviews (1997 and 2002). We defined three types of arts-related activity exposures: *lessons in musical arts, lessons in performing arts*, and *use of a musical instrument in the home*. Lessons in the arts exposures were defined by the CDS II parent interview questions with the primary caregiver (PCG) indicating the types of lessons in which the child was involved. We defined arts-related participation as taking lessons in one or more of the following: dance, drama, music-instrument, music-singing, music-not specified. We then combined these arts-related lessons into two categories in order to have adequate sample sizes for the outcome estimation procedures: lessons in dance or drama were grouped as performing arts, and the three types of music lessons were grouped as music.

Use of a musical instrument in the home was determined by the CDS II parent interview questions indicating whether there was an instrument in the family's home and the frequency of the instrument use by the child. We then defined the instrument exposure as three ordered categories: no instrument in the home, instrument in the home that is not used by the child, and instrument in the home that is used by the child (at least once per year).[2]

Covariates

The primary goal of the analysis plan was to remove potential confounding—that is, to make the assumption of exchangeability as plausible as possible. We accomplished this in two ways. First, we selected a more extensive set of covariates than has been used in prior research, described below. The literature on causal inference offers some guidance for selecting covariates (Wooldridge, 2009). In general, any potential confounders (variables that influence the outcome and the exposure) should be included. Implementing this principle, however, is less obvious than it might appear, especially when considering whether to add additional covariates above and beyond a core set that seems fairly obvious (e.g., family income). In considering additional covariates, the key issue is whether they are associated with the outcome and exposure within strata defined by the other covariates.

As discussed above, the choice of covariates in earlier research is limited, and some studies lack rather obvious confounders, like maternal education. Even holding all other aspects of the analyses constant, one would expect improved (or more plausible) estimates of the effect of the arts simply by using more comprehensive data. The PSID includes better measures of some covariates included in prior research. For example, earlier research generally includes some (crude) measure of parental SES. However, the PSID includes comprehensive measures of the family's income including child-specific expenditures, as well as measures of the family's assets, another indicator of economic resources. Other key potential confounders include parent IQ, parenting behaviors, and characteristics of the home as a learning environment. We selected covariates that capture the characteristics, behaviors and other factors that predicted participation in arts lessons after reviewing the prior research on arts involvement. This includes information about the child, the parents and the family/household, listed below.

Child. Child covariates included in the analyses are as follows: gender, number of siblings living with child, race/ethnicity, prior reading and math assessment scores, behavior problems, positive behaviors, ability self-concepts, low birthweight, repeated grade, involvement in other extracurricular activities, activity limitations, health conditions, vision or hearing impairments, and overall health status.

Parent and family. Parent and family covariates included in the analyses are as follows: family structure, IQ, years of education, work status, income, child-specific expenditures, assets, number of times child has changed schools, HOME scale score, family encourages hobbies, warmth towards child, communication with child, monitoring of child's activities, and involvement in child's school activities.

Outcomes

We examine two broad categories of outcomes: outcomes measured during adolescence in the CDS III, and outcomes measured during early adulthood. The CDS III outcomes we examined were as follows. The *Woodcock Johnson Revised Test of Achievement* measures child achievement in reading and math using three subtests (Letter-Word, Passage Comprehension, and Applied Problems). The Letter-Word and Passage Comprehension scores were also combined to make a broad reading score. The *Wechsler Intelligence Scales for Children (WISC) Digit Span* measures short-term working memory. The *Behavior Problems Index* measures the incidence and severity of child behavior problems including externalizing or aggressive behavior and internalizing or withdrawn behavior as reported by the PCG. The *Positive Behaviors Scale* measures positive aspects of children's lives including social competence, self-control, self-esteem, obedience, and persistence as reported by the PCG. *Ability Self-Concept* scales measure the child's perception of "self" in relation to their perceived ability in a particular domain. The three domains/subscales were Math, Reading, and Global (whole self).

A second set of analyses examined the effect of lessons in arts education on outcomes in early adulthood using the 2007 wave of the TA study: participation in arts activities; frequency of participation in arts activities (e.g., once a month, several times a week); high school completion; high school GPA; earnings (total annualized earnings in thousands of dollars for all jobs reported by participant that occurred during the prior calendar year [2006]).

Missing Data

The study maintained high rates of follow-up across the three waves of the CDS. The 1997 CDS-I successfully completed interviews with 2,394 families, providing information on 3,563 children. In 2002–2003, CDS recontacted families in CDS-I who remained active in the PSID panel as of 2001. CDS-II successfully reinterviewed 2,019 families (91%) who provided data on 2,907 children and adolescents aged 5 to 18 years. These 2,907 individuals are the individuals who contribute to our analyses. Some of these individuals did not participate in the first wave of the CDS (but had been eligible); others did not participate in the 2007–2008 data collection. Individuals who did not participate in CDS-II were not included in the analysis; these individuals lacked the key measures of arts participation. For that reason, we used the sampling weights from the CDS-II to make the data representative of the original 1997 population (which correct for any unrepresentativeness in the data involving key demographic characteristics due to attrition). Of the 2,907 individuals, 963 participated in the TA study; the remainder participated in the CDS-III. The two groups are the sample sizes for the analysis of child development and young adult outcomes, respectively. We handle individuals lost to follow-up and cases lacking information on required variables through multiple imputation as discussed below.

Structure of Analysis

Our analyses proceeded in a series of steps. The first step involved imputing missing values for incomplete observations. The second step involved estimating statistical models predicting exposure to arts education. The third step involved calculating the predicted probability of exposure (the propensity score) and the inverse probability of treatment (IPT) weights; finally, we conducted weighted analyses of the multiply imputed data.

Step 1. Address Missing Data Using Multiple Imputation

Multiple imputation has advantages over common methods for handling missing data in longitudinal studies, such as complete-case analysis. Analyses of multiply imputed (MI) data proceeds in two steps. The first involves filling in a statistical prediction for the missing data based on the values of variables for which data are available. The imputation model assumes the data are "missing at random," meaning that those that did and did not participate in the study can differ in terms of the observed covariates (Allison, 2002; Little & Rubin, 2002). However, conditional on those covariates, the values of those who participated must be representative of all individuals with that profile of covariates. Analysis of complete cases makes this same assumption. The advantage of MI is that it allows one to include observations, such as those individuals who participated in two of the three waves of data collection, which can improve the precision of parameter estimates.

The second step involves obtaining an estimate of key parameters from imputation-specific analyses using "Rubin's rules" (Little & Rubin, 2002; Schafer, 1997, 1999). One performs separate analyses for each of the imputed datasets, and then the imputation-specific parameter estimates are averaged together. One then calculates the variance for the overall estimate using the imputation-specific variances plus a "penalty" that reflects variation across the imputations. The latter reflects the fact that we do not observe the missing data (by definition); the penalty increases the standard error of the estimate to reflect this uncertainty. The key is that this increase is smaller than what occurs when incomplete cases are dropped from the analysis entirely. We performed each of the follow-up steps six times, once for each imputation.

Step 2. Estimate Models Predicting Lessons in the Arts and Calculate Propensity Scores

The propensity score is the predicted probability of exposure given a unit's profile of covariates. One can calculate the propensity score using any method that predicts exposure using the covariates. We performed logistic regression on a dichotomous indicator (yes-no) of arts participation, thereby predicting a child's "propensity" for arts participation. One typically assumes that this proclivity to participate (Y*) is a function of the determinants of characteristics, behaviors, and the costs of participation, some of which are unobserved. The latter is represented by $\in$, which is assumed distributed according to the negative extreme value distribution. One can write the probability of participation as follows:

$$\begin{aligned} P(Y_i=1). &= P(Y_i^*>0) \\ &= P(\in_i > -\underline{X}_i\,\underline{B}) \\ &= (1+\exp(\underline{X}_i\,\underline{B}))^{-1}, \end{aligned}$$

where $\underline{X}$ is an $N \times K$ matrix of the K covariates and $\underline{B}$ is the $K \times 1$ vector of corresponding coefficients. Logistic regression produces maximum-likelihood estimates of the latter.

Having estimated the coefficients of that model, one can generate the predicted probability of exposure using the model estimates. It is calculated using the above formula with $\underline{B}$ replaced by the parameter estimates. For music exposure, the propensity score represents the predicted probability of taking lessons or playing an instrument at home. We have a separate propensity score for the three exposures. They differ to the extent that each exposure relates to the covariates differently. For example, an individual observation might have a propensity score of 50%. Of the individuals with the same propensity score, some will actually participate and others will not. Of those with a propensity score of 50%, for example, half will participate and half will not. The essence of the method is to compare individuals with the same propensity score who do and do not participate. The intuition is that those individuals with the same propensity score are all else equal. This is the exchangeability assumption in this context, and it is important to remember that the propensity score depends on the same fundamental assumption as ordinary regression (i.e., no unobserved confounding, or ignorability). The principal advantage of the propensity score is that it summarizes the covariates effectively—that is, matching cases with the same propensity score is as effective in removing confounding as working with all the covariates that were used to create the propensity score. This convenience facilitates checking key model assumptions on which regression rests. *One would expect carefully done regression to produce the same results as analyses based on the propensity score.*

Step 3. Calculate IPT Weights

One can use the propensity score to adjust comparisons of two groups for confounders in several ways, such as matching. Our analyses involve a particularly intuitive form, the inverse probability of treatment (IPT) weights. The weights are a transformation of the propensity score (Cole & Hernán, 2008). The IPT weight is one over the probability that the exposure was actually experienced. For the arts participants, the IPT weight is one over the propensity score; for nonparticipants, it is one over one minus the propensity score. (One minus the propensity score is the probability of *not* being treated.) Analyses using these weights essentially correspond to analyses of a pseudo-population in which exposure and the covariates are unrelated. In other words, these weights allow one to use the sample data to represent a hypothetical population where there is no confounding—where the distribution of the covariates is the same for those who are and are not taking lessons in the arts.[3] When weighted, one can compare the two groups using simple statistical tests. This includes assessing key assumptions that are important for both regression and propensity score-based analyses but not readily apparent in the former, namely, checking covariate balance and support (Caliendo & Kopeinig, 2008).

Step 4. Conduct Weighted Analyses of Multiply Imputed Data

In the fourth step of the analysis, we conducted statistical analyses involving the IPT-weighted data. These tests represent comparisons of means for continuous outcomes and differences in the predicted probability of the outcome for categorical outcomes, such as participation in the arts in young adulthood.

Results

This section includes both descriptive and analytic results. First, we describe our analysis sample and their exposure to arts education after conducting multiple imputation, defined as lessons in musical arts or performing arts (dance or drama). We then discuss the predictors of participation in both types of lessons as the first step of the propensity score analyses. Finally, we present our main analytic results that examine the relationship between lessons in the arts and playing an instrument in the home and key developmental outcomes during childhood, adolescence, and young adulthood.

Describing the Sample

Our analyses involve all 2,907 children participating in the second wave of data collection, including those who did not participate in the third wave. Of these cases, 1,944 contribute at least some data to the outcome analysis for children ages 12 and under at the start of the study. A total of 963 individuals (ages 11 to 14 at study baseline) contributed at least some data to the analysis of young adult outcomes reported in the TA component of the study.

Analyses of the raw data confirm that listwise deletion would have reduced the sample size dramatically. The vast majority of observations in the analysis of the child development outcomes had missing data for one or more variables. To some extent, this reflects the large number of variables used in the analysis and features of the measures used. For example, a substantial portion (64%) of the sample lacked the Passage Comprehension measure we used as a covariate. Much of this missingness is accounted for by the child's age: Four of five observations lacking this measure were ineligible in the first wave because they were age six and younger. Nonetheless, about one in four older children were lacking this measure as well.

Tables 2a and 2b provide descriptive data on the multiply imputed data used for the analyses for those individuals ages 19 and under in 2007. Appendix Tables C.1 and C.2 in the online version of the journal focus on the data used in the analyses of outcomes in young adulthood (>19 in 2007). Each table provides descriptive data on the outcomes and then on the covariates. The tables disaggregate the data by participation in music lessons and performing arts lessons, respectively. Panel A in each table provides the means and standard deviations (*SD*) for those participating; Panel B, for those not participating. Panel C describes the sample as a whole. One can compare Panels A and B to see raw, unadjusted differences between those who do and do not participate. These associations reveal the potential confounders that might create spurious relationships between

Table 2a

Child Development Outcomes and Descriptive Statistics, by Lessons in Musical Arts (n = 1,944)

	Lessons in Musical Arts						
	A. Involved		B. Not		C. All		
	Mean	*SD*	Mean	*SD*	Mean	*SD*	*p* Value
Involved in Lessons in Musical Arts	100%	0	0%	0	22%	41%	NA
I. OUTCOMES (2007)							
Applied problems score	101.15	21.60	95.70	25.48	96.88	24.82	.00**
Behavior Problem Index score	6.66	5.91	7.71	6.52	7.48	6.41	.02*
Broad reading score	102.09	17.32	96.18	17.73	97.46	17.88	.00**
Digit span score	15.45	4.32	14.80	4.04	14.94	4.13	.05
Global self-concept score	4.14	0.64	4.06	0.67	4.08	0.66	.30
Letter-word score	103.79	16.80	98.28	17.32	99.48	17.44	.00**
Math self-concept score	4.76	1.08	4.81	1.05	4.80	1.06	.61
Passage comprehension score	99.91	17.46	94.51	18.29	95.68	18.29	.00**
Positive Behavior Index score	4.25	0.56	4.19	0.64	4.20	0.63	.27
Reading self-concept score	5.14	1.03	5.06	1.08	5.07	1.07	.20
II. COVARIATES (1997)							
Gender (male = 1; female = 0)	44%		52%		51%		.00**
Number of siblings living with child	1.41	1.30	1.49	1.35	1.47	1.34	.17
Race/ethnicity (reference category: white non-Hispanic)							
Black-NH	32%		43%		41%		.00**
Hispanic	9%		7%		7%		.03*
Other race/eth.	3%		3%		3%		.58
Baseline Measures of Child Development Outcomes							
Broad reading score	104.66	17.91	99.45	16.76	100.58	17.19	.00**
Broad math score	103.71	17.89	99.60	18.80	100.49	18.70	.00**

(continued)

Table 2a **(continued)**

	Lessons in Musical Arts						
	A. Involved		B. Not		C. All		
	Mean	*SD*	Mean	*SD*	Mean	*SD*	*p* Value
Behavior Problem Index score	8.87	6.36	11.41	8.24	10.86	7.92	.00**
Positive Behavior Index score	3.76	1.49	3.17	1.88	3.30	1.82	.00**
Ability Self Concepts: Reading	1.39	2.35	1.95	2.55	1.83	2.52	.00**
Ability Self Concepts: Math	1.34	2.27	1.85	2.44	1.74	2.41	.00**
Global Self Concept	1.50	2.54	2.11	2.75	1.98	2.72	.00**
Other Child Characteristics							
Low birthweight	7%		8%		8%		.37
Repeated grade	8%		13%		12%		.00**
Child participates in sports	32%		31%		32%		.58
Child participates in community groups	33%		21%		24%		.00**
Activity limitations	5%		8%		7%		.01*
Chronic cond.-mental	5%		8%		7%		.01*
Chronic cond.-physical	22%		24%		23%		.34
Chronic cond.-other	14%		19%		18%		.01*
Vision or hearing imp.	7%		7%		7%		.74
PCG reported child health (1-5 scale)	1.53	0.74	1.68	0.83	1.65	0.81	.00**
Family and Parental Demographics							
Family Structure (reference category: Two-parent family)							
Bio parent- no partner	19%		32%		29%		.00**
Bio parent- w/ partner	3%		5%		4%		.17
IQ score- HH	9.49	2.28	8.97	2.26	9.08	2.27	.00**
Years of education- HH	13.56	2.39	12.70	2.18	12.89	2.25	.00**

(continued)

Table 2a **(continued)**

	Lessons in Musical Arts						
	A. Involved		B. Not		C. All		
	Mean	*SD*	Mean	*SD*	Mean	*SD*	*p* Value
Currently employed- HH	87%		78%		80%		.00**
Average of family income between 1997 & 2002 in ten-thousands	0.43	0.50	0.29	0.40	0.32	0.43	.00**
Child-specific expenditures in thousands	3.54	2.74	2.99	2.61	3.11	2.65	.00**
Total family assets in ten-thousands	0.15	0.60	0.08	0.38	0.10	0.44	.00**
Times child changed schools past year	0.05	0.25	0.08	0.34	0.07	0.32	.07
Other Family and Parental Characteristics							
HOME scale score	1.22	0.19	1.09	0.20	1.12	0.21	.00**
Family encourages hobbies	90%		83%		84%		.00**
Parental warmth	3.99	0.59	3.87	0.68	3.90	0.66	.00**
How often PCG talks w. child about school	5.38	1.12	5.01	1.52	5.09	1.45	.00**
How often PCG talks w. child about interests	4.02	0.88	3.90	0.98	3.93	0.96	.02*
How often PCG talks about books/reads w. child	3.19	1.27	3.05	1.34	3.08	1.33	.00**
No. child's books (1-5 scale)	4.75	0.62	4.48	0.89	4.54	0.85	.00**
No. child's friends PCG knows (1-5 scale)	3.90	1.13	3.59	1.24	3.65	1.23	.00**
No. times PCG volunteered at child's school[a]	0.68	1.17	0.52	0.98	0.55	1.03	.00**
No. times PCG attended child's school events[a]	1.39	1.03	0.98	1.07	1.07	1.08	

Note. Means are calculated using multiple imputations. Standard deviations (*SD*s) are calculated using the first imputation for continuous variables only. HH = head of household; PCG = primary caregiver.
[a]Log-transformed.
*p < .05. **p < .01.

Table 2b

Child Development Outcomes and Descriptive Statistics, by Lessons in Performing Arts (*n* = 1,944)

	Lessons in Performing Arts						
	A. Involved		B. Not		C. All		
	Mean	*SD*	Mean	*SD*	Mean	*SD*	*p* Value
Involved in Lessons in Performing Arts	100%	0	0%	0	12%	32%	NA
OUTCOMES (2007)							
Applied problems score	102.03	19.09	96.20	25.43	96.88	24.82	.02*
Behavior Problem Index score	6.74	5.99	7.58	6.46	7.48	6.41	.12
Broad reading score	101.69	16.01	96.91	18.01	97.46	17.88	.01*
Digit span score	15.07	3.99	14.93	4.14	14.94	4.13	.71
Global self-concept score	4.23	0.55	4.06	0.67	4.08	0.66	.07
Letter-word score	102.69	16.42	99.05	17.50	99.48	17.44	.04*
Math self-concept score	4.66	1.05	4.82	1.06	4.80	1.06	.09
Passage comprehension score	99.93	15.60	95.12	18.53	95.68	18.29	.01*
Positive Behavior Index score	4.24	0.53	4.19	0.64	4.20	0.63	.28
Reading self-concept score	5.19	1.06	5.06	1.07	5.07	1.07	.20
COVARIATES (1997)							
Gender (male = 1; female = 0)	19%		55%		51%		.00**
Number of siblings living with child	1.49	1.40	1.47	1.33	1.47	1.34	.81
Race/ethnicity (reference category: white non-Hispanic)							
Black-NH	35%		42%		41%		.01*
Hispanic	5%		8%		7%		.05
Other race/eth.	5%		3%		3%		.04*
Baseline Measures of Child Development Outcomes							
Broad reading score	102.39	17.22	100.35	17.18	100.58	17.19	.13
Broad math score	101.12	16.72	100.41	18.94	100.49	18.70	.66

(continued)

Table 2b **(continued)**

	Lessons in Performing Arts						
	A. Involved		B. Not		C. All		
	Mean	*SD*	Mean	*SD*	Mean	*SD*	*p* Value
Behavior Problem Index score	10.31	6.35	10.93	8.05	10.86	7.92	.61
Positive Behavior Index score	3.14	1.94	3.31	1.80	3.30	1.82	.11
Ability Self Concepts: Reading	1.66	2.52	1.85	2.52	1.83	2.52	.20
Ability Self Concepts: Math	1.48	2.29	1.78	2.43	1.74	2.41	.04*
Global Self Concept	1.75	2.67	2.01	2.73	1.98	2.72	.10
Other Child Characteristics							
Low birthweight	8%		8%		8%		.91
Repeated grade	8%		12%		12%		.02*
Child participates in sports	32%		32%		32%		.88
Child participates in community groups	36%		22%		24%		.00**
Activity limitations	4%		8%		7%		.02
Chronic cond.-mental	6%		7%		7%		.40
Chronic cond.-physical	24%		23%		23%		.92
Chronic cond.-other	17%		18%		18%		.58
Vision or hearing imp.	6%		7%		7%		.57
PCG reported child health (1-5 scale)	1.51	0.74	1.67	0.82	1.65	0.81	.00**
Family and Parental Demographics							
Family Structure (reference category: Both bio parents)							
Bio parent- no partner	26%		29%		29%		.22
Bio parent- w/ partner	4%		4%		4%		.47
IQ score- HH	9.43	2.21	9.03	2.28	9.08	2.27	.00**
Years of education- HH	13.76	2.16	12.77	2.24	12.89	2.25	.00**

(continued)

Table 2b **(continued)**

	Lessons in Performing Arts						
	A. Involved		B. Not		C. All		
	Mean	*SD*	Mean	*SD*	Mean	*SD*	*p* Value
Currently employed- HH	86%		80%		80%		.01*
Average of family income between 1997 & 2002 in ten-thousands	0.38	0.27	0.32	0.44	0.32	0.43	.01*
Child-specific expenditures in thousands	3.61	2.60	3.04	2.65	3.11	2.65	.00**
Total family assets in ten-thousands	0.19	0.76	0.09	0.38	0.10	0.44	.00**
Times child changed schools past year	0.06	0.29	0.08	0.32	0.07	0.32	.26
Other Family and Parental Characteristics							
HOME scale score	1.21	0.18	1.11	0.21	1.12	0.21	.00**
Family encourages hobbies	92%		83%		84%		.00**
Parental warmth	4.04	0.58	3.88	0.67	3.90	0.66	.00**
How often PCG talks w. child about school	5.38	1.16	5.05	1.48	5.09	1.45	.00**
How often PCG talks w. child about interests	4.09	0.87	3.90	0.97	3.93	0.96	.00**
How often PCG talks about books/reads w. child	3.37	1.27	3.04	1.33	3.08	1.33	.00**
No. child's books (1-5 scale)	4.85	0.47	4.50	0.88	4.54	0.85	.00**
No. child's friends PCG knows (1-5 scale)	3.96	1.09	3.61	1.24	3.65	1.23	.00**
No. times PCG volunteered at child's school	0.86	1.25	0.51	0.99	0.55	1.03	.00**
No. times PCG attended child's school events	1.28	1.03	1.04	1.08	1.07	1.08	.00**

Note. Means are calculated using multiple imputations. Standard deviations (*SD*s) are calculated using the first imputation for continuous variables only. HH = head of household; PCG = primary caregiver.

*$p < .05$. **$p < .01$.

arts participation and outcomes. The last column provides the *p*-value for the between-group comparison.

Participation in Music Lessons and Child Development Outcomes

First looking at Table 2a, one can see that 22% of the sample of individuals used in the analysis of child development outcomes (Section I of Table 2a) participated in music lessons. The first section provides information on the key outcomes for these children, differentiating them according to whether they participated in music lessons. These figures are quite consistent with prior research: One sees many associations between music participation and improved outcomes. For 5 of the 10 outcomes, children in music lessons scored better, for example, higher on the Letter-Word score and lower on the behavior problems index. One can see, however, that some outcomes were not (even) associated with participation (e.g., global self-concept). In terms of practical magnitude, the differences involved are rather small. Figure 2 provides some sense of the practical significance of the coefficients by presenting effect sizes. For example, on the Applied Problems score (the first outcome listed), participants and nonparticipants differed by 5.45 points (=101.15-95.70). The SD in the nonparticipant group is 25.48, implying an effect size of .21 (=5.45/25.48), the height of the left-most bar in the figure.[4]

In the second panel of the table (Section II), similar data are presented for the covariates identified as potential confounders. One can see that many of these are associated with participation in music lessons. For example, 44% of participants are male compared with 52% of nonparticipants. Simple comparisons of participants and nonparticipants of course reflect both gender differences and the effect of music participation per se. This section also provides information on 1997 characteristics of the child. Many of these are earlier measurements of the variables treated as outcomes in 2007. Children engaged in music lessons differed from other children at other time points. Consider behavior problems—children participating in music lessons in 2002 scored 2.55 points lower on this measure in 1997 (i.e., before music lessons were observed). This difference is actually larger than that observed in 2007—after music participation was measured. *Baseline characteristics predict future music lessons participation even when the reverse is not true* (i.e., music lessons do not predict future achievement). Global self-concept predicts participation, but participation does not predict the same concept subsequently. All of these associations are very modest.

Family demographics also strongly predict participation. Children from single-parent families are underrepresented among participants. Nineteen percent of children taking lessons live in single-parent families; this is substantially less than that for nonparticipants (32%; $p < .01$). In particular, the detailed household financial variables including the amount of family

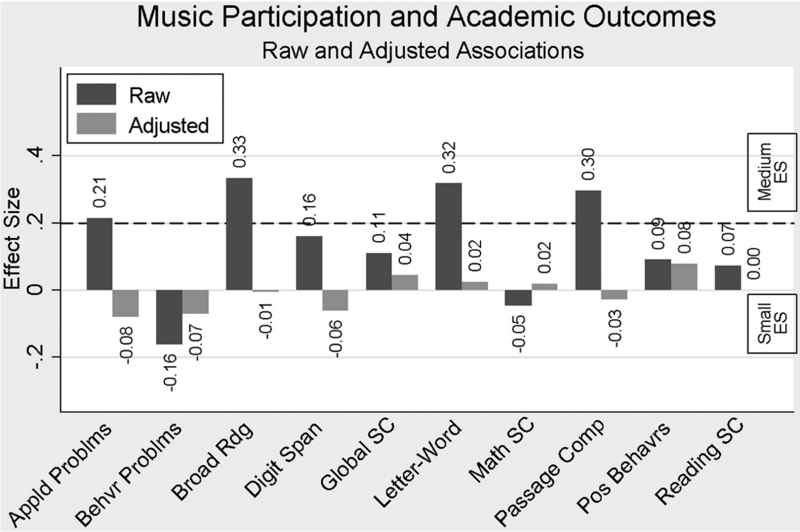

Figure 2. **Raw and propensity score adjusted effect sizes for childhood music lessons on cognitive, social, and emotional outcomes during adolescence (*n* = 1,944).**

Source. Panel Study of Income Dynamics.

Note. Outcomes are as follows (starting at the left side of the figure): *Woodcock Johnson Revised Test of Achievement* (WJ) *Applied Problems* subtest (math skills); *Behavior Problems Index* (child behavior problems including externalizing or aggressive behavior and internalizing or withdrawn behavior); *Broad Reading score* (combined Letter-Word and Passage Comprehension scores); *Wechsler Intelligence Scales for Children (WISC) Digit Span* (short-term working memory); *Global Self-Concept* (selfesteem); *WJ Letter Word* subtest (vocabulary/reading skills); *Math Self-Concept* (child's perception of 'self' in relation to their perceived ability in math); *WJ Passage Comprehension* (reading skills); *Positive Behaviors* (social competence, self-control, self-esteem, obedience, and persistence); *Reading Self-Concept* (child's perception of 'self' in relation to their perceived reading ability). Effect sizes considered small or medium based on Cohen (1988).

assets and total child-related expenditures predicted participation in music lessons. *In general, every family characteristic points toward the social advantages of children participating in music lessons.*

Participation in Performing Arts Lessons and Child Development Outcomes

Table 2b differs from Table 2a in that figures are disaggregated by participation in lessons in performance arts rather than music lessons. Twelve percent of the sample participated in such lessons. The pattern across outcomes is very similar to that for music participation. In terms of the predictors of participation, perhaps most striking is the increased gender imbalance. Only 19% of those participating in lessons in the performing arts are male. While not surprising, these and other anticipated relationships confirm the quality of our measures of participation.

This section of the table also shows that baseline measures of the outcomes do not predict lessons in performance arts. The pattern across other child characteristics is also very similar. For example, there appears to be no differences between children who participate in performing arts lessons and children who do not participate in terms of birthweight (low or normal), chronic health conditions, and participation in sports. However, some key child characteristics such as parent reported child health, participation in community groups, and having repeated a grade predicted performing arts participation. As with the music participation figures shown in Table 1, every family demographic characteristic (e.g., assets, income, parent years of education) is positively associated with children's participation in performing arts lessons.

Predictors of Participation in the TA Sample

Appendices C.1 and C.2 in the online version of the journal present similar figures for the smaller sample of older children participating in the TA survey. We provide these figures for completeness, but one would anticipate that the relationships would be very similar. The two sets of figures should only differ to the degree that the predictors of participation changed as children age. They are in fact very similar. As with the child development sample data shown in Tables 2a and 2b, the children's baseline outcome scores and other characteristics were not predictive of the early adulthood outcomes. Likewise, many of the family characteristics and parenting behavior measures were positively associated with participation in both music and performing arts lessons for the young adulthood sample.

Calculating the Propensity Scores

Appendix D in the online version of the journal presents the results of the logistic regression used to calculate the propensity scores. Coefficient estimates are presented as marginal effects—the effect of the covariate on the predicted probability of participation holding the other covariates constant at sample means. These estimates show that a range of factors are significant, independent predictors of participation. As happens with multiple regression, these adjusted differences can be the reverse of unadjusted differences. For example, African American children are 7.3 percentage points less likely to participate in music lessons in the unadjusted sample. Adjusting for the other variables, African American children are 10.5 percentage points *more likely* to participate, all else equal. This reversal of direction demonstrates the strength of the other covariates in predicting participation and of their covariation with race and ethnicity. In particular, the unadjusted racial difference reflects the overrepresentation of children of color in strata (defined by SES and other characteristics) where participation is lower. Within those strata, however, children of color appear more likely to participate.

Of the socioeconomic measures, family income is the most strongly predictive. That relationship is curvilinear—income increases the likelihood of music participation but the effect of added income drops at higher levels of income. The parents of children who participate in both music and performing arts also have higher IQ scores, more years of education and family assets, and also tend to spend more on their children. Some of the other family and parental characteristics predict participation as well. The Home Observation for Measurement of the Environment Inventory (HOME) score is a very strong predictor of participation in both music and performing arts lessons. The HOME score is frequently used in child development research as an overall assessment of the quality of cognitive stimulation and the emotional support that the child receives from the family, measured using both parent-report and observational items.

Thus, children whose families are more supportive of their child and encourage learning activities at home are more likely to be involved in other learning activities, such as arts education. Furthermore, several other indicators of parental involvement in the child's daily life (e.g., attending school events, knowing child's friends) also predicted participation. Not surprisingly, having a musical instrument in the home was a strong predictor of participation in music lessons; children who have a musical instrument in the home were 18 percentage points more likely to have participated in music lessons.

Several of the child's baseline measures on the cognitive outcomes are significant. Some of these are not significant, and one explanation is multicollinearity among the measures—because they are correlated, distinguishing the distinct contribution of each measure is difficult. In some types of analyses, this issue would represent a problem—for example, if one were truly interested in whether reading skills was more predictive of participation in arts education than, say, math skills. However, in the case at hand, we are most interested in obtaining the best estimate of the effect of participation itself and want to remove any confounding by other variables, regardless of whether the confounding is created by math or reading skills.

Outcome Analyses

The "Effect" of Lessons in Musical Arts on Child Development Outcomes

Table 3a shows the propensity score weighted estimate of the effect of lessons on key outcomes observed in the 2007 wave of the Child Development Survey. None of the effects are statistically significant—none

Table 3a

Effects of Arts Participation on Child Development Outcomes (n = 1,944)

		Letter-Word	Passage Comprehension	Broad Reading	Applied Problems	Digit Span	Behavior Problems	Positive Behaviors	Reading Self-Concept	Math Self-Concept	Global Self-Concept
Lessons in Musical Arts[a]	β	0.42	–0.53	–0.10	–2.03	–0.25	–0.46	0.05	0.00	0.02	0.03
	SE	1.60	2.50	1.70	2.68	0.52	0.50	–0.25	0.08	0.11	0.08
Lessons in Performing Arts[a]	β	0.73	2.08	1.68	1.70	0.35	0.03	0.03	0.11	–0.14	0.03
	SE	2.61	2.89	2.63	4.37	0.54	0.68	0.07	0.14	0.15	0.10
Instrument in Home[b]											
No instrument in home	β	–3.18	–1.16	–2.34	–0.75	–0.57	–3.18	–1.16	–2.34	–0.75	–0.57
	SE	2.38	2.03	1.95	2.77	0.74	2.38	2.03	1.95	2.77	0.74
Use instrument in home	β	–1.68	0.57	–0.64	–0.42	–0.20	–1.68	0.57	–0.64	–0.42	–0.20
	SE	2.13	1.78	1.77	3.24	0.49	2.13	1.78	1.77	3.24	0.49

Note. None of the results were significant at conventional levels of statistical significance, and for that reason, columns indicating significance are omitted.

[a]Reference category: no lessons.

[b]Reference category: instrument in home but not played by child.

Table 3b
Effects of Arts Participation on Young Adult Outcomes (*n* = 963)

		HS Graduation[b]	HS GPA	Adult Participation	Earnings
Lessons in Musical Arts[a]	β	0.06	0.10	0.22*	–0.11
	SE	0.48	0.08	0.05	0.15
Lessons in Performing Arts[a]	β	See text	–0.09	0.26*	0.28
	SE		0.13	0.13	0.19

[a]Reference category: no lessons.
[b]The reported statistic is the marginal effect of the participation on the probability of high-school graduation.
*$p < .05$.

are more than 1.65 standard errors from zero. The lack of statistical significance reflects both the small practical magnitude of the relationships as well as the loss of statistical precision. The latter refers to the fact that the standard errors are larger than those in unadjusted analyses. Incorporating the propensity score weights inflated standard errors generally by about 50%. *All in all, assuming exchangeability and including the broader array of covariates, lessons in musical arts do not influence future child development outcomes.*

Returning to Figure 2, one can see the importance of adjustment for confounding—the figure shows the effect size for the adjusted and unadjusted between-group differences. Two things are striking about the figure. First, the unadjusted differences were consistent yet small. Cohen (1988) defined effect sizes of no larger than .20 *SD*s as "small"; otherwise less than .50 as "medium". Secondly, there remain no differences between musical arts participants and nonparticipants after adjusting for observed confounding (i.e., using the IPT weights).

The "Effect" of Lessons in Performance Arts on Child Development Outcomes

As with musical arts, the data (Table 3) show no relationship between performance arts and child development outcomes. The effects are significant in neither the practical nor the statistical sense.

The "Effect" of Lessons in Musical Arts on Outcomes in Young Adulthood

Table 3b presents the effect of arts lessons on outcomes in early adulthood. Note that high school graduation and participation in the arts are dichotomous outcomes, and the data present the marginal effect calculated using logistic regression. The coefficient represents the effect of music lessons on the probability of completing high school and of participating in the arts. Results indicate that children who participate in music lessons in mid-childhood are approximately 22 percentage points more likely to participate in the arts in young adulthood. The effect of participation is sizable for high school graduation (5 percentage points) but is not statistically significant. The lack of statistical significance reflects the large standard error of the estimate.

The "Effect" of Lessons in Performance Arts on Outcomes in Young Adulthood

The results in Table 3b indicate that children who participate in performing arts lessons in mid-childhood are 26 percentage points more likely to participate in the arts in young adulthood. They are also more likely to participate more frequently.

The effect on high school graduation is not reported because 100% of those individuals in the data who participated in lessons in performance arts finished high school. When we examine this descriptively, however, the weighted percentage of those finishing high school is 10 percentage points higher among those students taking performing arts ($p < .01$).

The "Effect" of Using an Instrument in the Home on Child Development Outcomes

In supplemental analyses, we considered an added exposure, "playing an instrument in the home." To reiterate, we defined the instrument exposure as three ordered categories: no instrument in the home, instrument in the home that is not used by the child, and instrument in the home that is used by the child (at least once per year). The coefficients in Table 3a would therefore be interpreted with respect to the reference category, having an instrument in the home that is not played by the child. However, none of the effects for these models were significant, meaning that there are no significant differences in late childhood schooling and behavioral outcomes between children who experienced different levels of the instrument exposure in mid-childhood.

Discussion

Using nationally representative data, our analyses provide little support for the notion that the well-documented associations between arts participation and development in childhood and beyond reflect the effect of the former per se. These results suggest that the prior research in this area apparently fail to account for the range of confounding factors that influence both child outcomes and arts participation. A strength of data like the PSID is that it includes a broad array of key family measures, including multiple measures of the family's resources, information about parenting (e.g., parent discusses interests with child, parent knows child's friends, etc.) and relatively objective assessments of the home learning environment (i.e., HOME score). These measures are strong predictors of child outcomes in the child development literature at large. Including them here has protected our results from confounding, and consequently, eliminated any associations between arts participation and a broad array of cognitive and behavioral outcomes.

As noted, earlier studies include rather crude measures of SES. Our analyses incorporate a multiyear comprehensive measure of family income reported at the time lessons in the arts or playing an instrument in the home was recorded. Of the socioeconomic measures, family income was the most strongly predictive of children's participation in arts education. The parents of children who participate in both music and performing arts also have higher IQ scores, more years of education and family assets and also tend to spend more on their children. Indeed, enrolling one's child in arts lessons characterize "concerted cultivation"; parenting styles of middle- and upper-income parents that involve organized activities and other structured investments in children's development (Lareau, 2003). As noted, after adjusting for family factors, we found no associations between participating in arts lessons and children's development. These findings are consistent with some of the work in this area that shows that once family SES is included, measures of concerted cultivation activities do not strongly influence child development (Bodovski & Farkas, 2008; Cheadle, 2008).

We also found that children's characteristics confound the effects of arts participation. Prior cognitive outcomes and performance predict arts participation and subsequent outcomes. The strength of these associations is troubling. *Arts participation may be yet another advantage enjoyed by these children. If arts participation produces better outcomes, it builds on the child's prior success*. In this way, arts education may reinforce inequality generated by child- and family-level advantages.

Some of the research on the effects of music education suggest that music participation improves children's self-esteem, self-efficacy, and overall self-perception, and these positive attitudes can increase the student's motivation for all curricular and extracurricular endeavors, such as studying (Hallam, 2010). For these reasons, we included ability and global self-concept measures in our analyses as both predictor variables and also as outcomes themselves in the CDS III. Looking at the descriptive statistics in Tables 2a and 2b, one can see that while arts participants have *higher* self-concept scores (reading, math, global) compared to nonparticipants in 2007, they had *lower* scores in these measures in the preexposure period, 1997, suggesting that participants had improved in these measures over time. Our analyses tested whether this relationship was indeed *caused* by lessons in arts education. However, we did not find any significant adjusted relationships between these measures and arts participation.

On the other hand, we did find that when children participate in arts education they are more likely to be engaged in the arts as young adults. This is an important finding because engagement in some type of activity or interest in young adulthood is associated with performance, persistence, and overall well-being across life stages (Fredricks & Eccles, 2006; Ryan & Deci, 2000). Indeed, Lerner (2003) notes that "a young person may be said to be *thriving*, then, if he or she is involved across time in such healthy, positive relations with his or her community … marked by making culturally valued contributions to self, others, and institutions" (p. 173). Though arts education may not influence GPA, our findings do support the idea that engagement in the arts begets engagement in the arts, and this is a culturally valued activity indicative of well-being in adulthood.

Relatedly, some research has documented benefits to those participating in *group* arts activities. This type of experience builds social confidence, social networks, self-esteem, and sense of belonging (Eccles & Barber, 1999; Eccles et al., 2003; Hallam & Prince, 2000; Sward, 1989). We did not have that type of detail of the arts exposure in this study, but this certainly can be explored in future research. However, this more refined measurement of arts exposure would further complicate causal inference as discussed below.

Note that our analyses incorporate preferred methodology for dealing with the challenges inherent to this type of research. One of these includes the use of propensity scores rather than ordinary regression. However, it is important to recognize that propensity scores and regression both rest on the same key assumption—exchangeability. Furthermore, many of the problems with regression could be addressed through better empirical practice, including (a) checking the support of the covariates to make sure they overlap sufficiently, (b) checking functional form, and (c) checking for outliers. A regression analysis that attended to these issues carefully would likely produce results like those reported here. In that light, the key contribution of this article may be simple: We analyzed data that offered a rich set of covariates and we used them.

The principal limitation of our analysis is the possibility of unobserved confounding. Such confounding does not always produce associations that are larger than the true causal effect. In our case, the lack of significant effects—when substantial associations exist—still may reflect unobserved confounding. In this case, one would have to believe that children who participate in music lessons are *more* disadvantaged than other children *within the strata defined by the covariates* (i.e., all else equal). That possibility seems implausible in light of how strongly higher income and other advantages predict involvement. Nonetheless, such confounding is possible.

Another explanation for a lack of adjusted (causal) relationships is that reports of music or performing arts participation or outcomes are not accurate. It is reassuring, then, to see that our results document the associations that are present in other research. These associations suggest that key constructs are measured accurately.

Another possibility is that the exposure to the arts is defined too broadly. Children taking music lessons include many children that may neither persist in those lessons over time nor be truly engaged more generally (e.g., practicing very little). The subset of children who actually do so may benefit, and the inclusion of these low participants may "water down" the true benefits of higher participation. The issue of how broadly to define an exposure is a common challenge in causal inference. For example, whether children participate in lessons (at all) reflects the choice to start lessons (only). Children who participate over time are distinguished by a decision about whether to continue. As a behavior is defined more narrowly, those who engage in that behavior are increasingly self-selected; the more narrow the behavior or exposure, the more choices are involved. Each of these decisions provides an opportunity for unobserved confounding to occur. Finding a group of children who are comparable to those who participate in arts lessons is difficult; finding a group comparable to those who participate a great deal is more challenging still.

Though our study included a comprehensive set of child development outcomes, other unexamined benefits of arts lessons may still exist. For example, education leaders from Organization of Economic Cooperation and Development (OECD) countries have recently shown an interest in understanding the "collateral" benefits of arts education in promoting economic development. They argue that the arts foster innovation and entrepreneurship through the opportunity to think creatively and work collaboratively without the existence of a right or wrong answer (Winner, Goldstein, & Vincent-Lancrin, 2013). Because innovation is viewed as central to economic growth, they consider the cultivation of these skills as central to a 21st-century education. While these undocumented relationships remain a possibility, innovation and entrepreneurship were not measured in our study. Furthermore, these factors are difficult to measure and separate from other outcomes (like those included in our study).

Implications for Future Research and Data Collection

These analyses highlight the value of good measurement and large samples found in large, national datasets and the importance of capturing of possible confounders in future research. Our analyses incorporate a multiyear comprehensive measure of family income reported at the time that children were exposed to lessons in the arts or were playing an instrument in the home. As noted, other studies employ a rather limited range of covariates, and as a result, these studies likely confound the effect of arts education with other advantages these children enjoy. Furthermore, many of the studies of arts education involve small selective samples.

On the other hand, large national studies address many topics, and the depth of measurement for any specific topic may be lacking. As discussed above, such is the case with the definition of arts education in the PSID. The measure of arts exposure in this dataset lacks, for example, any indication of the level or extent of participation. While understanding the effects of a more extensive or refined definition of arts education raises causal challenges, the starting point for such analyses is understanding these processes. To do so, one needs measures differentiating children by their level of participation. Analyses of such data would require samples large enough to provide meaningful numbers of children at each level of participation. Only large, national datasets are likely to provide such numbers. Furthermore, our analyses suggest that children who do and do not participate in the arts differ in many ways, thus longitudinal data are essential. One can see that past measures of achievement predict arts participation, and it is likely that children at different levels of involvement are as differentiated. The relationship between achievement and music education and exposure likely unfolds over time.

In that sense, the PSID represents a promising start, and the benefits of adding indicators of arts education are apparent. The shallowness of that measurement, however, limits the potential for further exploration of these data. As the participants now have aged into early adulthood, the window for measuring arts involvement has closed. However, other studies of young children might include better and more detailed measures of arts participation. Studies such as the Early Childhood Longitudinal Study offer the potential for further, future explorations of the link between arts education and children's outcomes. Because some emerging work has found benefits of arts education on the oral language development for students with limited English proficiency (Chappell & Cahnmann-Taylor, 2013; Greenfader et al., 2015), future research should consider this potentially important subpopulation. Analyses of data that capture student's entrepreneurial and innovative abilities, suggested by the OECD as a potential benefit of arts education, will also be relevant to the policy discussions surrounding public investments in arts education.

Conclusion

Southgate and Roscigno (2009) report that 93% of Americans believe that a well-rounded education includes music education. Our analyses need not weaken public support for music education. What it does do is weaken efforts to bolster support for arts education by linking it to benefits outside of the arts, including test scores. Such efforts are well intentioned, and if such benefits existed, it would move discussions away from the inherent value of the arts and toward more "tangible" benefits. What these analyses suggest, however, is that public discussions of the value of the arts cannot avoid the issue of "values" per se through a linkage to test scores and measures of a child's achievement. The arts have been embedded in the human experience since our first societies, and they will remain to be embedded in our life and culture regardless of their impact on student learning (Winner et al., 2013).

A range of studies reveal an association between quality of life and the arts, and other studies reveal substantial public willingness to pay for the arts. Indeed, we found that students who participated in the arts as a child or adolescent were more likely to participate in the arts as a young adult. These figures imply support for arts education—the availability of dancers and musicians in the future depends on their training now. If policy makers believe that such activities have broader social benefits, then lessons for children and youth today represent an investment in the nation's cultural capital.

Notes

We wish to recognize the National Endowment for the Arts *Art Works* grant program for funding this work through contract order number C11-119. We would like to thank Jacquelynne Eccles, George Farkas, Greg Duncan, Sue Hallam, Paul Hanselman, and several members of the NEA Research and Analysis office for comments on previous drafts.

[1]As discussed in Foster (2010), many papers include a standard disclaimer in the discussion, reminding that the article reveals an association only. In these instances, authors never explain why such an association is of much interest and proceed to give their findings a causal interpretation. For example, Wetter et al. (2009) summarize their findings as "Musical training evidently correlates with children's better performance at school, but is obviously part of a multifactorial dependence. Continuous musical training appears to help achieve and maintain school performance at a high level over time." If the first statement is correct, then their results cannot be interpreted as implying the second.

[2]While the CDS included arts-related interview questions in all three waves, the question about child participation in extracurricular activities or lessons in the arts in the first wave could not support our definition of arts exposure. This is because the question did not have the parent identify whether the activity in which the child participated fell under the category of arts, sports, or community organizations. We could not explicitly identify arts exposure and therefore did not examine the effects of arts exposure in the early- to mid-childhood time period.

[3]This is equivalent to sampling weights commonly used in studies to make a sample representative of a population (Thompson, 1992). Individuals who are underrepresented in a group are "weighted up" (their experience is counted multiple times in calculating sample characteristics).

[4]Another way of expressing this is that an effect size of .20 is equivalent to moving an individual on a scale from the 50th percentile to the 57th percentile. (This 7-point increase corresponds to moving from the mean of a normally distributed variable to a z score of 1.20.) An effect size of .5 would correspond to a movement from the mean to the 69th percentile.

References

Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: SAGE Publications.

Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. Z. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology*, *83*, 111–130.

Barber, B. L., Eccles, J. S., & Stone, M. R. (2001). Whatever happened to the Jock, the Brain, and the Princess? Young adult pathways linked to adolescent activity involvement and social identity. *Journal of Adolescent Research*, *16*(5), 429–455.

Becker, G. S., & Tomes, N. (1994). Human capital and the rise and fall of families. In G. S. Becker (Ed.), *Human capital: A theoretical and empirical analysis with special reference to education* (3rd ed., pp. 257–298). Chicago: University of Chicago Press.

Bilhartz, T. D., Bruhn, R. A., & Olson, J. E. (1999). The effect of early music training on child cognitive development. *Journal of Applied Developmental Psychology*, *20*(4), 615–636.

Bobek, D., Zaff, J., Li, Y., & Lerner, R. M. (2009). Cognitive, emotional, and behavioral components of civic action: Towards an integrated measure of civic engagement. *Journal of Applied Developmental Psychology*, *30*(5), 615–627.

Bodovski, K., & Farkas, G. (2008). "Concerted cultivation" and unequal achievement in elementary school. *Social Science Research*, *37*(3), 903–919.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22*(1), 31–72.

Catterall, J., Chapleau, R., & Iwanaga, J. (1999). *Involvement in the arts and human development: General involvement and intensive involvement in music and theater arts. Champions of change: The impact of the arts on learning*. Los Angeles: Graduate School of Education & Information Studies, University of California at Los Angeles.

Chan, A. S., Ho, Y.C., & Cheung, M.-C. (1998). Music training improves verbal memory. *Nature*, *396*(6707), 128.

Chappell, S. V., & Cahnmann-Taylor, M. (2013). No child left with crayons: The imperative of arts-based education and research with language "minority" and other minoritized communities. *Review of Research in Education*, *37*(1), 243–268.

Cheadle, J. E. (2008). Educational investment, family context, and children's math and reading growth from kindergarten through the third grade. *Sociology of Education*, *81*(1), 1–31.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cole, S. R., & Hernán, M. A. (2008). Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology*, *168*(6), 656–664.

Corenblum, B., & Marshall, E. (1998). The band played on: Predicting students' intentions to continue studying music. *Journal of Research in Music Education*, *46*(1), 128–140.

Costa-Giomi, E. (1999). The effects of three years of piano instruction on children's cognitive development. *Journal of Research in Music Education*, *47*(3), 198–212.

Deci, E. L., & Ryan, R. M. (2008). Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychology/Psychologie Canadienne*, *49*(1), 14–23.

Eccles, J. S., & Barber, B. L. (1999). Student council, volunteering, basketball, or marching band: What kind of extracurricular involvement matters? *Journal of Adolescent Research*, *14*(1), 10–43.

Eccles, J. S., Barber, B. L., Stone, M., & Hunt, J. (2003). Extracurricular activities and adolescent development. *Journal of Social Issues*, *59*(4), 865–889.

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*(2), 175–194.

Elpus, K., & Abril, C. R. (2011). High school music ensemble students in the United States: A demographic profile. *Journal of Research in Music Education*, *59*(2), 128–145.

Feldman, A. F., & Matjasko, J. L. (2007). Profiles and portfolios of adolescent school-based extracurricular activity participation. *Journal of Adolescence*, *30*, 313–332.

Fisher, D. (2001). Early language learning with and without music. *Reading Horizons*, *42*(1), 8.

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*(1), 73–84. doi:10.1177/002242940605400106

Forgeard, M., Winner, E., Norton, A., & Schlaug, G. (2008). Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning. *PLoS ONE*, *3*(10), e3566.

Foster, E. M. (2010). Causal inference and developmental psychology. *Developmental Psychology*, *46*(6), 1454–1480. doi:10.1037/a0020204

Fredricks, J. A., Alfeld-Liro, C. J., Hruda, L. Z., Eccles, J. S., Patrick, H., & Ryan, A. M. (2002). A qualitative exploration of adolescents and commitment to athletics and the arts. *Journal of Adolescent Research*, *17*(1), 68–97. doi:10.1177/0743558402171004

Fredricks, J. A., & Eccles, J. S. (2006). Is extracurricular participation associated with beneficial outcomes? Concurrent and longitudinal relations. *Developmental Psychology*, *42*(4), 698.

Gardiner, M. E., Fox, A., Knowles, F., & Jeffrey, D. (1996). Learning improved by arts training. *Nature*, *381*(6580), 284.

Greenfader, C. M., Brouillette, L., & Farkas, G. (2015). Effect of a performing arts program on the oral language skills of young English learners. *Reading Research Quarterly*, *50*(2), 185–203. doi:10.1002/rrq.90

Hallam, S. (2010). The power of music: Its impact on the intellectual, social and personal development of children and young people. *International Journal of Music Education*, *28*, 269.

Hallam, S., & Prince, V. (2000). *Research into instrumental music services*. London: DfEE.

Helmrich, B. H. (2010). Window of opportunity? Adolescence, music, and algebra. *Journal of Adolescent Research*, *25*(4), 557–577. doi:10.1177/0743558410366594

Hetland, L. (2000, Autumn–Winter). Learning to make music enhances spatial reasoning. *Journal of Aesthetic Education*, *34*, 179–238.

Hetland, L., & Winner, E. (2001). The arts and academic achievement: What the evidence shows. *Arts Education Policy Review*, *102*(5), 3–6.

Ho, Y.-C., Cheung, M.-C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology*, *17*(3), 439–450.

Hofferth, S., Davis-Kean, P. E., Davis, J., & Finkelstein, J. (1998). *The child development supplement to the Panel Study of Income Dynamics: 1997 user guide*. Ann Arbor: Survey Research Center, Institute for Social Research, University of Michigan.

Hurwitz, I., Wolff, P., Bortnick, B., & Kokas, K. (1975). Nonmusical effects of the kodaly music curriculum in primary grade children. *Journal of Learning Disabilities*, *8*(3), 167–174.

Kinney, D. W. (2010). Selected nonmusic predictors of urban students: Decisions to enroll and persist in middle school band programs. *Journal of Research in Music Education*, *57*(4), 334–350. doi:10.1177/0022429409350086

Klinedinst, R. E. (1991). Predicting performance achievement and retention of fifth-grade instrumental students. *Journal of Research in Music Education*, *39*(3), 225–238.

Lareau, A. (2003). *Unequal childhoods: Race, class and family life*. Berkeley: University of California Press.

Larson, R. W. (2000). Toward a psychology of positive youth development. *American Psychologist*, *55*(1), 170–183.

Lerner, R. M., Dowling, E., & Anderson, P. M. (2003). Positive youth development: Thriving as the basis of personhood and civil society. *Applied Developmental Science*, *7*(3), 172–180.

Lerner, R. M., Lerner, J. V., Almerigi, J. B., Theokas, C., Phelps, E., Gestsdottir, S., . . .von Eye, A. (2005). Positive youth development, participation in community youth development programs, and community contributions of fifth-grade adolescents: Findings from the first wave of the 4-H Study of Positive Youth Development. *Journal of Early Adolescence*, *25*(1), 17–71. doi:10.1177/0272431604272461

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Hoboken, NJ: John Wiley and Sons.

Magne, C., Schon, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioural and electrophysiological approaches. *Journal of Cognitive Neuroscience*, *18*, 199–211.

Mahoney, J. L., Cairns, B. D., & Farmer, T. W. (2003). Promoting interpersonal competence and educational success through extracurricular activity participation. *Journal of Educational Psychology*, *95*(2), 409.

Mehr, S. A., Schachner, A., Katz, R. C., & Spelke, E. S. (2013). Two randomized trials provide no consistent evidence for nonmusical cognitive benefits of brief preschool music enrichment. *PLoS ONE*, *8*(12), e82007. doi:10.1371/journal.pone.0082007

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science*, *22*(11), 1425–1433.

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, *19*(3), 712–723.

Morrison, S. J. (1994). Music students and academic growth. *Music Educators Journal*, *81*(2), 33–36.

Munte, T. F., Nager, W., Beiss, T., Schroeder, C., & Erne, S. N. (2003). Specialization of the specialized: Electrophysiological investigations in professional musicians. In G. Avanzini, C. Faienza, D. Minciacchi, L. Lopez, & M. Majno (Eds.), *The neurosciences and music* (pp. 112–117). New York: Academy of Sciences.

Parsad, B., & Spiegelman, M. (2012). *Arts education in public elementary and secondary schools: 1999–2000 and 2009–10*. Retrieved from http://files.eric.ed.gov/fulltext/ED530715.pdf

Patel, A. D. (2009). Music and the brain: Three links to language. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (pp. 208–216). New York: Oxford University Press.

Piro, J. M., & Ortiz, C. (2009). The effect of piano lessons on the vocabulary and verbal sequencing skills of primary grade students. *Psychology of Music*, *37*(3), 325–347.

Rickard, N. S., Bambrick, C. J., & Gill, A. (2012). Absence of widespread psychosocial and cognitive effects of school-based music instruction in 10-13-year-old students. *International Journal of Music Education*, *30*(1), 1–22.
Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67.
Scales, P., Benson, P., & Roehlkepartain, E. (2011). Adolescent thriving: The role of sparks, relationships, and empowerment. *Journal of Youth and Adolescence*, *40*(3), 263–277.
Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman and Hall.
Schafer, J. L. (1999). Multiple imputation: A primer. *Statistical Methods in Medical Research*, *8*(1), 3–15.
Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science*, *15*(8), 511–514.
Schellenberg, E. G. (2006). Long-term positive associations between music lessons and IQ. *Journal of Educational Psychology*, *98*(2), 457.
Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. *British Journal of Psychology*, *102*(3), 283–302. doi:10.1111/j.2044-8295.2010.02000.x
Schellenberg, E. G., & Moreno, S. (2010). Music lessons, pitch processing, and g. *Psychology of Music*, *38*(2), 209–221. doi:10.1177/0305735609339473
Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child's brain and cognitive development. *Annals of the New York Academy of Sciences*, *1060*(1), 219–230.
Schneider, T. W., & Klotz, J. (2000). *The impact of music education and athletic participation on academic achievement*. Paper presented at the annual meeting of the Mid-South Educational Research Association, Bowling Green, KY.
Southgate, D. E., & Roscigno, V. J. (2009). The impact of music on childhood and adolescent achievement. *Social Science Quarterly*, *90*(1), 4–21.
Sward, R. (1989, Winter). Band is a family. *Today's Music Educator*, 26–27.
Thompson, S. K. (1992). *Sampling*. New York: Wiley.
Vaughn, K. (2000). Music and mathematics: Modest support for the oft-claimed relationship. *Journal of Aesthetic Education*, *34*(3/4), 149–166.
Wetter, O. E., Koerner, F., & Schwaninger, A. (2009). Does musical training improve school performance? *Instructional Science*, *37*(4), 365–374.
Winner, E., & Cooper, M. (2000). Mute those claims: No evidence (yet) for a causal link between arts study and academic achievement. *Journal of Aesthetic Education*, *34*(3/4), 11–75. doi:10.2307/3333637
Winner, E., Goldstein, T., & Vincent-Lancrin, S. (2013). *Art for art's sake? The impact of arts education*. Educational Research and Innovation, OECD Publishing. Retrieved from http://www.oecd.org/edu/ceri/arts.htm
Winsler, A., Gara, T., Alegrado, A., Castro, S., & Tavassolie, T. (2016). *Selection into, and academic benefits from, arts-related elective courses in middle-school among low-income, ethnically diverse youth*. Washington, DC: National Endowment for the Arts.
Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*(4), 420–422.
Wooldridge, J. M. (2009). *Introductory econometrics: A modern approach* (4th ed.). Mason, OH: South-Western, Cengage Learning.

Zulauf, M. (1993). Three-year experiment in extended music teaching in Switzerland: The different effects observed in a group of French-speaking pupils. *Bulletin of the Council for Research in Music Education*, *119*, 111–121.

Manuscript received Month XX, XXXX
Final version received Month XX, XXXX
Accepted Month XX, XXXX