Dev Psychol. Author manuscript.

Musical Instrument Engagement in Adolescence Predicts Verbal Ability Four Years Later: A Twin and Adoption Study

Daniel E. Gustavson[1,2], **Naomi P. Friedman**[3,4], **Michael C. Stallings**[3,4], **Chandra A. Reynolds**[5], **Hilary Coon**[6], **Robin P. Corley**[4], **John K. Hewitt**[3], **Reyna L. Gordon**[2,7,8]

[1]Department of Medicine, Vanderbilt University Medical Center, Nashville, TN

[2]Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN

[3]Institute for Behavioral Genetics, University of Colorado Boulder, Boulder, CO

[4]Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO

[5]Department of Psychology, University of California Riverside, Riverside, CA

[6]Department of Psychiatry, University of Utah, Salt Lake City, UT

[7]Department of Otolaryngology – Head & Neck Surgery, Vanderbilt University Medical Center, Nashville, TN

[8]Department of Psychology, Vanderbilt University, Nashville, TN

Abstract

Individual differences in music traits are heritable and correlated with the development of cognitive and communication skills, but little is known about whether diverse modes of music engagement (e.g., playing instruments vs. singing) reflect similar underlying genetic/environmental influences. Moreover, the biological etiology underlying the relationship between musicality and childhood language development is poorly understood. Here we explored genetic and environmental associations between music engagement and verbal ability in the Colorado Adoption/Twin Study of Lifespan behavioral development and cognitive aging (CATSLife) project. N=1684 adolescents completed measures of music engagement and intelligence at approximately age 12 and/or multiple tests of verbal ability at age 16. Structural equation models revealed that instrument engagement was highly heritable (a^2=.78), with moderate heritabilities for singing (a^2=.43) and dance engagement (a^2=.66). Adolescent self-reported instrument engagement (but not singing or dance engagement) was genetically correlated with age 12 verbal intelligence, and still was associated with age 16 verbal ability even when controlling for age 12 full-scale intelligence, providing evidence for a longitudinal relationship between music engagement and language beyond shared general cognitive processes. Together, these novel findings suggest that shared genetic influences in part accounts for phenotypic associations between music engagement and language, but there may also be some (weak) direct benefits of music engagement on later language abilities.

Correspondence: Correspondence concerning this article should be addressed to Daniel Gustavson, Department of Medicine, Vanderbilt University Medical Center, 2215 Garland Ave, 511H Light Hall, Nashville, TN, 37232. daniel.e.gustavson@vumc.org.

Keywords

heritability; twin study; adoption study; musicality; vocabulary; language

Introduction

Language and music are uniquely human traits that allow us to communicate a wide range of information. An extensive literature shows significant sharing of sensory, cognitive, and neural resources between these domains (Kraus & White-Schwoch, 2016), and there are robust associations between individual differences in musical and language abilities throughout the lifespan (Mankel & Bidelman, 2018; Politimou, Dalla Bella, Farrugia, & Franco, 2019). Findings of enhanced speech perception sensitivity in musicians versus non-musicians initially suggested a relationship between musical expertise and speech/language abilities (Bidelman & Alain, 2015; Magne, Schön, & Besson, 2006; Marie, Magne, & Besson, 2011). This relationship is not only auditory/sensory in nature; it also characterizes associations between musical traits and higher language structure (i.e. complex syntax and reading: Brod & Opitz, 2012; Gordon, Jacobs, Schuele, & McAuley, 2015b; Swaminathan, Schellenberg, & Venkatesan, 2018).

The biology underlying music-language associations has primarily been studied with cognitive neuroscience methods. It is not yet known whether a common genetic basis drives the phenotypic correlations in addition to an overlap of brain networks (Patel, 2014). Individual predispositions to learning music may also account for enhanced language abilities (Schellenberg, 2015, 2019), rooted in the proposition that musically-relevant aural sensitivity to the acoustic and syntactic structure of speech bootstraps early language acquisition (Brandt, Gebrian, & Slevc, 2012). In the present study, we evaluate the relationship between musical engagement and language ability in a genetically informative developmental sample. Genetic studies provide an opportunity to explore the biological drivers of human complex traits such as musicality and can inform theoretical models by testing whether results are more consistent with causality or shared genetic influences.

The concept of “musicality” includes music perception abilities, musical behaviors, music engagement, music training, and more (Honing, 2018). Population-based survey studies have highlighted robust individual differences for musicality (Mas-Herrero et al., 2013; Mullensiefen, Gingras, Musil, & Stewart, 2014), which correspond well to objective assessments (Law & Zentner, 2012). Large-scale phenotyping efforts (i.e., Goldsmiths Musical Sophistication Index) suggest that measures of interest, lessons, skills, and total number of instruments played are highly interrelated (Mullensiefen et al., 2014).

The neural development of musicality over the course of childhood/adolescence may have special importance for language development (Zatorre, 2013; Zuk & Gaab, 2018). A multitude of cognitive neuroscience studies (Zhao & Kuhl, 2016) suggest that children use “musical hearing” (sensitivity to acoustic and structural features shared between music and speech) to develop their musical and language abilities in an intertwined and parallel manner (Brandt et al., 2012) via shared neural mechanisms (de Diego-Balaguer, Martinez-Alvarez, & Pons, 2016). Indeed, music aptitude in children is associated with a wide array of developmental language traits including phonological awareness (Ozernov-Palchik, Wolf, & Patel, 2018; Woodruff Carr et al., 2014), vocabulary (Piro & Ortiz, 2009), second language learning (Slevc & Miyake, 2006), and spoken grammar (Gordon et al., 2015c; Swaminathan & Schellenberg, 2019). Associations between music and language traits are evident even from preschool age (Politimou et al., 2019; Sallat & Jentschke, 2015) and persist into adolescence and adulthood (Dittinger et al., 2016). There is also potential clinical relevance of atypical music skills as a risk factor for developmental speech and language disorders (see Atypical Rhythm Risk Hypothesis in Ladanyi et al., 2020). This body of work has led researchers to postulate a strong biological link between music and language processes (Fitch, 2017; Honing, 2018).

Considered cumulatively, an emerging theoretical framework suggests that associations between musicality and language skills could be driven by genetic correlations (Ladanyi et al., 2020; Schellenberg, 2019). First, family-based approaches (primarily twin studies) have shown moderate heritability of musicality across objective and subjective measures, including music perception, achievement, interest, and practice habits in adolescents and adults (Coon & Carey, 1989; Hambrick & Tucker-Drob, 2015; Mosing et al., 2014a; Ullén et al., 2014; Vinkhuyzen, van der Sluis, Posthuma, & Boomsma, 2009). Careful exploration of the heritability of musicality and traits related to it with twin modeling is important groundwork for future work that could elucidate the molecular basis of these processes (Gingras et al., 2015). Importantly, existing genetic studies on musicality that have focused on intelligence quotient (IQ) and nonverbal abilities suggest that associations between music and cognitive abilities are explained primarily by genetic influences (Mosing, Pedersen, Madison, & Ullen, 2014b). Performance on verbal tasks (such as vocabulary and verbal fluency) are also moderately heritable (e.g., Gustavson et al., 2019), so their relationships with music engagement could also be genetically mediated. Second, there is only limited meta-analytic evidence that music lessons enhance language skills in children and young adults (Gordon, Fehd, & McCandliss, 2015a; Sala & Gobet, 2017). Indeed, a substantial body of work that initially interpreted superior performance on speech/language tasks as a transfer from musical training to language expertise (e.g., Patel, 2014) is now being revisited as evidence of self-selection for music training (Bidelman & Mankel, 2019; Schellenberg, 2019).

These patterns of findings are highly consistent with genetic correlations, wherein individuals with genetic predispositions for stronger language or cognitive abilities are more likely to seek out musical training (i.e., self-selection or pleiotropy). In other words, phenotypic associations between music engagement and cognition may reflect a common set of underlying genetic influences rather than causal associations, but direct evidence is needed to test these hypothesized genetic associations. Genetic correlations would be consistent with evidence for associations between music and language traits even in individuals with little to no music training. Investigation of neural plasticity effects of music experience associated with the development of language expertise (see Zatorre, 2013; Zuk & Gaab, 2018) lead to the possibility of shared genetic variation common to musicality and language. One could hypothesize that the biological function of such a set of genes subserves the development and activity of auditory-motor-language brain networks. An open question is thus whether effects of music training on neural and behavioral correlates of speech/language learning can be directly attributed to music training, driven by factors such as self-selection (including genetic predisposition), or both (Schellenberg, 2015). Twin and family studies can shed light on these possibilities by examining whether phenotypic correlations are explained by genetic and/or environmental correlations.

Moreover, longitudinal studies can test for evidence supporting causality by examining whether musical engagement predicts individual differences in later language abilities controlling for concurrent general cognitive ability. If music engagement predicts later language abilities even after controlling for measures of IQ (or specific verbal IQ subtests), this pattern would be consistent with a beneficial effect of music on language development. Although these genetic and longitudinal methods cannot prove causation, they can potentially reject causal explanations (e.g., if longitudinal associations are entirely mediated by concurrent IQ). They may also provide converging evidence in addition to music training studies regarding whether these associations are driven by genetic pleiotropy (i.e., genetic confounding), direct phenotypic effects, or a combination of both.

The Current Study

The current study examined the genetic and environmental structure of adolescent self-reported music engagement measures and their associations with verbal abilities, using data from the Colorado Adoption/Twin Study of Lifespan behavioral development and cognitive aging (CATSLife) project (Wadsworth et al., 2019). This project includes both twin and family samples, allowing us to collectively and comprehensively leverage data from twins, biological siblings, and adopted siblings.

First, we examined associations among measures of engagement (e.g., interests, lessons, talent) within the domains of musical instrument, singing, and dance engagement at the age 12 assessment, including estimating their heritability and environmental influences at the latent construct level. The abovementioned groundwork suggests that various measures of musicality tap a set of related constructs that are explained, at least in part, by genetic predispositions. We hypothesize that various measures of music engagement (e.g., talent at playing instruments, total number of instruments played) will be correlated with one another and explained by a common set of genetic influences.

We also examined whether measures of musical instrument playing, singing, and dance exhibit similar genetic and environmental influences, or if they are explained by unique genetic and environmental factors. While evidence suggests possible differences between singing and other aspects of musicality (Slater et al., 2017), less is known about the genetic architecture of singing and how it relates to verbal skills. Vocalists are often collapsed with instrumentalists in existing phenotypic studies (Theorell et al., 2015), with some evidence for subtle differences in their relation to sensory and cognitive processes (Slater et al., 2017), but dancers have rarely been included in these studies despite evidence that musicality centrally involves the motor system (Patel & Iversen, 2014). As a basis for exploration of associations with language (or other relevant cognitive and psychosocial outcomes), it is crucial to quantify the shared and/or distinct genetic influences across these different aspects of music engagement. Such knowledge would be informative in light of phenotypic relationships between singing and language abilities (Clement et al., 2015; Hutchins, 2018). Given the importance of language competence in academic and life success (Bashir & Scavuzzo, 1992), evidence of musicality as a familial factor that can predict language/verbal outcomes would have significant public health impact (Ladanyi et al., 2020).

Finally, we tested the hypothesis that music engagement factor(s) would be correlated with a Verbal Ability latent factor assessed about 4 years later. In these growth-focused analyses, we controlled for age 12 IQ, which would support the idea that music engagement influences later verbal abilities controlling for baseline verbal and nonverbal abilities. We also examined whether the associations were most consistent with genetic and/or environmental correlations.

Methods

All studies were approved by the University of Colorado Boulder Institutional Review Board (Protocol Number: 14–0421; Title: Colorado Adoption Project/Twin Study of Lifespan behavioral development & cognitive aging [CATSLife]). Informed consent was obtained from each participant at each assessment.

Participants

Analyses were based on a total of 1684 individuals (830 females, 854 males) from the CATSLife project (Wadsworth et al., 2019). CATSLife comprises two samples with similar measures: the Colorado Longitudinal Twin Sample (LTS; Corley et al., 2019) and the Colorado Adoption Project (CAP; Wadsworth et al., 2019). LTS subjects included same-sex twin pairs (232 full monozygotic [MZ] twin pairs, 203 full dizygotic [DZ] twin pairs, and 2 unpaired twins) and CAP subjects included 177 biological sibling pairs, 176 adoptive sibling pairs, and 106 unpaired individuals. Participants completed questionnaire measures of music engagement and intelligence at about age 12 (*n*=1373, *M*=12.45 years, *SD*=0.38). They also completed verbal ability tests at about age 16 (*n*=1641, *M*=16.61 years, *SD*=1.20)[1]. Most participants completed both waves of assessment (*n*=1287). Data were included for all individuals with even partial data because they are still informative for the latent factor structure and heritability.

The LTS sample was recruited through the Colorado Department of Health based on twins born between 1984 and 1990. Twins are representative of the population of Colorado at that time (Corley et al., 2019; Rhea, Gross, Haberstick, & Corley, 2013). Participants identified as white (91.9%), Hawaiian or Pacific islander (0.2%), Asian (0.2%), American Indian or Alaskan (1.1%), more than one race (5.4%), or unknown or unreported race (1.1%). Hispanic individuals comprised 9.1% of the sample.

---

[1] The age range of LTS twins and CAP probands was small at the “age 12” (11.3 to 14.2 years) and “age 16” assessments (16.0 to 20.0 years). This was similar for CAP siblings at “age 12” (11.5 to 14.0 years; average absolute age difference from proband *M*=0.42 years, Range=0 to 1.17) but wider at “age 16” (15.83 to 36.1 years; average absolute age difference from proband *M*=0.91 years, Range=0 to 20.08) because this assessment included additional siblings not previously assessed longitudinally. Post-hoc analyses removing all 20 siblings who were older than 20.0 years at the “age 16” assessment had no impact on the phenotypic results, so they were included in all genetic analyses.

The CAP sample began recruitment in 1975 with the support of Denver social services agencies and area hospitals (Plomin & DeFries, 1983; Rhea, Bricker, Wadsworth, & Corley, 2013). Parents were recruited with a one-to-one ratio of adoptive and nonadoptive parents. Initially, the first younger sibling in the family was also enrolled, but later studies expanded to include other siblings. For these analyses, we chose siblings to pair with each proband based on the order of enrollment (i.e., first male or female sibling). If they had both a male and female sibling, we chose the sibling with the most data (i.e., completed both assessments). If both had complete data, we chose the same-sex sibling. Participants identified as white (92.0%), Asian (4.6%), American Indian or Alaskan (1.5%), more than one race (0.9%), or unknown or unreported race (0.6%). Hispanic individuals comprised 0.9% of the sample.

Measures

Age 12 Music Engagement.—For musical instrument, singing, and dance engagement, adolescents were asked how interested they were in the activity, whether they received formal instruction (yes/no), and how talented they were in that domain (0–3 scale from “no talent” to “really good”). The interest measure included 3 responses (“hate it”, “like it”, “love it”) plus a fourth response (“never tried it”). The final response was excluded from analyses because it was difficult to gauge the child’s interest if they had never tried the activity, and preliminary analyses indicated that similar results were observed if this option was included as the lowest option. All questionnaire responses were treated as ordinal variables in all analyses.

Finally, adolescents were asked to report the total number of musical instruments they have played, including singing. Because this item included both singing and instrument engagement, it was allowed to have a factor loading on both the Musical Instrument and Singing engagement latent factors reported in the results. Additionally, because responses were skewed, the item was binned into 3 categories (0 instruments, 1 instrument, and 2 or more instruments). Correlations with other study measures were nearly identical if continuous scores were used instead, but binning is preferred to obtain unbiased parameter estimates in genetic analyses (Derks, Dolan, & Boomsma, 2004). Although these items were not based on extant questionnaires, the items are similar to validated self-report measures such as the Goldsmiths Musical Sophistication Index (e.g., “I can play _____ musical instruments”, “I have never been complimented for my talents as a musical performer”, etc.), which has moderate to high reliability and were validated against objective listening tasks (Mullensiefen et al., 2014). Assessing music engagement and musical sophistication with as little as a single questionnaire item is widespread (Zhang & Schubert, 2019), so integrating across 3–4 measures from 3 music engagement domains represents a stepping-stone for future larger-scale genetic investigations, including task-based measures of musical ability (Zentner & Gingras, 2019).

Age 12 intelligence.—Intelligence was assessed with the Wechsler Intelligence Scale for Children-Revised Edition in CAP (WISC-R; Wechsler, 1974) and the 3rd edition in LTS (WISC-III; Kaufman, 1994). Our primary analyses focused on the full-scale IQ from both versions of the WISC. Some additional analyses were also conducted using the verbal IQ and performance IQ measures of the WISC separately. Verbal IQ subtests include vocabulary, information, comprehension, arithmetic, and similarities. Performance IQ subtests include picture completion, picture arrangement, block design, object assembly, and coding.

Age 16 verbal ability.—Verbal ability was assessed using the vocabulary subtest from the Wechsler Adult Intelligence Scale-Revised Edition in CAP (WAIS-R; Wechsler, 1981) or 3rd edition in LTS (WAIS-III; Wechsler, 1997), as well as 3 tests (vocabulary, letter fluency and category fluency) from the Specific Cognitive Abilities (SCA) battery that was developed for CAP based on earlier work from the Hawaii Family Study of Cognition (Defries, Plomin, Vandenberg, & Kuse, 1981) and completed in both subsamples. In the SCA vocabulary test, participants had 3 minutes to answer up to 50 multiple-choice items (part 1) or 4 minutes to answer up to 25 more difficult items (part 2). The total score was the average across parts 1 and 2. Letter fluency was the average score across two subtests (*S-P* and *G-T*) where participants had 3 minutes to name as many words as possible beginning with one letter and ending with another (e.g., beginning with an *S* and ending with a *P*). Category fluency was also the average score across two subtests where participants had 3 minutes to generate as many names of “things that are often round” and “things that are often metal” as possible. Earlier work using the LTS data demonstrated substantial genetic overlap among vocabulary and fluency measures (Gustavson et al., 2019).

Data Analysis

Analyses were conducted using Mplus (Muthén & Muthén, 1998–2017), which accounts for missing observations using pairwise deletion (weighted least squares, mean and variance adjusted, theta parameterization). Model fit was determined based on chi-square tests (χ^2), the root mean error of approximation (RMSEA), and the Comparative Fit Index (CFI). Models were considered to have good fit if they had χ^2 values less than twice the degrees of freedom, RMSEA values less than .06, and CFI values greater than .95 (Hu & Bentler, 1998). Significance of individual parameter estimates was established using standard error-based 95% confidence intervals (CIs) for phenotypic analyses and bootstrapped 95% CIs for genetic analyses (with 1000 iterations), both of which were confirmed with χ^2 difference tests ($p < .05$). Standard-error based confidence intervals are not appropriate for genetic models in which parameters are transformed (e.g., squaring genetic paths to determine heritability), bounded (e.g., heritability is bounded 0 to 1), and prone to be drawn from highly asymmetric distributions (Neale & Miller, 1997). It is therefore necessary to use bootstrapping or likelihood-based confidence intervals, only the former of which are available in Mplus.

Phenotypic analyses controlled for sex, race (white vs. non-white), ethnicity (Hispanic vs. non-Hispanic), and sibling type (MZ vs. DZ in LTS; biological vs. adoptive siblings in CAP) by regressing these measures on all study variables in the model. Phenotypic analyses also accounted for the clustering by family in the data with Mplus’s “type=complex” command, which yields standard errors and χ^2 statistics adjusted for non-independence of twin data (Rebollo, de Moor, Dolan, & Boomsma, 2006; Satorra & Bentler, 2001). Sex was the only variable systematically associated with study measures but controlling for it did not alter the patterns of results, even in preliminary genetic analyses (see Supplement Figure S1). Therefore, to simplify the models and aid in model convergence, genetic analyses presented here do not adjust for any covariates.

Genetically informed models were based on standard assumptions in twin and family designs (Neale & Cardon, 1992), in which the variance of a phenotype can be separated into proportions attributable to additive genetic influences (A), common environmental influences (C), and nonshared environmental influences (E). Additive genetic influences (A) are correlated at 1.0 in MZ twin pairs, 0.5 in DZ twin pairs and biological siblings, and 0.0 in adoptive siblings because MZ twins share 100% of their alleles identical-by-descent, DZ twins and biological siblings share, on average, 50% of their alleles identical-by-descent, and adoptive siblings are genetically unrelated. Common/shared environmental influences (C) correlate 1.0 for all twins and siblings because they are environmental factors that make siblings in a family more similar to one another. Nonshared environmental influences (E) do not correlate in any of the groups by definition. We also assume equal means and variances within pairs and across sibling types. These standard assumptions for univariate analyses extend to the multivariate analyses described here, including situations where phenotypic correlations are decomposed into their genetic (r_g), shared environmental (r_c), and nonshared environmental components (r_e). Finally, by incorporating both twin and biological/adoptive sibling data, we assume that same-aged siblings' (i.e., twins) E influences are equivalent to the E influences for different-aged siblings (i.e., biological or adoptive siblings).

Model fitting approach.—First, to ensure the self-reported music engagement measures fit well within a single domain, we tested a common pathway model of instrument engagement data only. Second, we added measures of singing and dance engagement to the model to describe the genetic and environmental overlap among latent factors for music instrument, singing, and dance engagement. To facilitate convergence for this model, we fixed all residual ACE paths for music engagement measures that were estimated at 0 to 0 (including singing and dance measures that were estimated at 0 in the initial model).

Next, in phenotypic correlational and regression analyses, we estimated associations between the music engagement factor(s) from the first step with the age 12 IQ and an age 16 Verbal Ability latent variable. When fitting this model, we first ensured that the common pathway model of the age 16 verbal ability measures fit well and tested for evidence that means and residual variances of certain measures needed to be freed across twin and biological/adoptive sibling subsamples (particularly the WISC and WAIS tests where participants completed different versions). Indeed, model fit was substantially improved by freeing the means and variances for age 12 IQ and age 16 WAIS vocabulary. In models with measures that did not have identical variances, standardized estimates are presented separately for each sample (covariances estimates were still constrained to be equal across sample, which did not impact model fit, but the total correlations differ because of the freed residual variances). See supplement Figure S2 for the multivariate model of the age 16 Verbal Ability alone.

Finally, to the extent we observed phenotypic associations between Musical Instrument engagement, full-scale IQ, and Verbal Ability factors, we decomposed these associations into their genetic and environmental components using genetic models. In these analyses we fit a bivariate Cholesky decomposition and computed genetic (A), shared environmental (C), and nonshared environmental (E) correlations from the output. We were also interested in testing a direction of causation model (Heath et al., 1993), which compares the model fit of the bivariate Cholesky with a model in which music engagement directly predicts later verbal ability. This model can potentially reject a causal model if replacing the ACE cross-paths from the Cholesky with a single direct path from music engagement to later verbal ability results in a significantly worse fit than the bivariate Cholesky (though it cannot prove the causal model if the competing models fit similarly). In other words, if one trait causes the other, then the ACE factors affecting music engagement only indirectly influence the second trait verbal ability through the direct phenotypic path. However, this model is only adequately powered when the two traits substantially differ in modes of inheritance (e.g., substantially different proportions of genetic and environmental effects; see Heath et al. 1993, Table 4).

Power.—Analyses were based on a secondary data source collected about 20 years ago. The original data collection effort was not designed with this specific analysis in mind, so all subjects with available data were included and no a priori power calculation was conducted. Post-hoc power calculations (Soper, 2018) confirmed that analyses were well-powered to detect small phenotypic correlations among music engagement latent factors and their potential correlations with IQ and verbal ability. Specifically, to achieve 80% power detect correlations of r=.10 with 14 measured variables, 4 latent variables, α=.05, the estimated minimum N was 1,454. This calculation did not consider the non-independence of data nested within families, which should have little effect on this phenotypic power (and nesting within families was controlled for in the actual analyses).

Results

Descriptive Statistics

Descriptive statistics for all measures are displayed in Table 1, and phenotypic correlations among all measures are displayed in the supplement (Tables S1 and S2). Cross-sibling cross-trait correlations are also displayed in the supplement (Tables S3 and S4).

Genetic and Environmental Model of Musical Instrument Engagement

The multivariate genetic model of the musical instrument engagement measures alone (Figure 1) fit the data acceptably, χ^2(158)=303.09, p<.001, RMSEA=.068, CFI=.972. In this model and all others presented here, percent variance explained in individual measures (shown in rectangles) by their respective latent factors can be computed by squaring the factor loadings (shown on single-headed arrows) from the latent variables (shown in ovals). Variance in latent factors explained by genetic (A), shared environmental (C), and nonshared environmental influences (E) can also be computed by squaring their factor loadings.

As shown in Figure 1, most of the variance in the four music engagement measures was explained by a Music Engagement latent factor. Genetic influences (a^2) accounted for 77.8% of the variance in this latent factor, 95% CI [.55, .94]. The remaining variance was explained evenly between shared environmental influences, c^2=.11, 95% CI [.00, .29], and nonshared environmental influences, e^2=.11, 95% CI [.03, .21]. Of the remaining variance in the individual music engagement indicators not already explained by the latent factor, most was explained by nonshared environmental influences (which include measurement error). These results suggest that the Music Engagement latent factor captured nearly all the genetic influences on the individual indicators.

Genetic and Environmental Associations Among Instrument, Singing, and Dancing Engagement

Figure 2 displays the genetic correlations among the Musical Instrument, Singing, and Dance engagement latent factors, which also demonstrated good fit, $\chi^2(869)=1179.36$, $p<.001$, RMSEA=.042, CFI=.967. The Music Instrument engagement factor was only weakly-to-moderately associated with the factors for Singing, r=.19, 95% CI [.11, .27], and Dance engagement, r=.23, 95% CI [.16, .30] (see Table 2 for phenotypic correlations adjusted for covariates). These correlations appeared to be driven by genetic correlations, but the genetic correlations were not statistically significant for Singing, r_g=.30, 95% CI [−.04, .60], nor for Dance engagement, r_g=.23, 95% CI [−.06, .51]. In contrast, the Singing and Dance engagement factors were strongly correlated with one another, r=.56, 95% CI [.50, .62]. Genetic influences accounted for almost all of this association, r_g=.88, 95% CI [.54, 1.0], explaining 86% of the phenotypic correlation. The shared environmental, r_c=.12, 95% CI [−1.0, 1.0], and nonshared environmental correlations, r_e=.20, 95% CI [−.05, .49], were not significant. The heritability of Singing engagement was a^2=.43, 95% CI [.18, .63], and the heritability of Dance engagement was a^2=.66, 95% CI [.38, .80]. Because the three music engagement domains were not strongly correlated with one another, we did not test higher-order common factor models.

This model included three residual correlations, informed by preliminary phenotypic analyses. We initially tested phenotypic associations between all measures of the same type (e.g., instrument lessons with singing lessons and dance lessons), but only 2 correlations were statistically significant: one between instrument lessons and dance lessons, and one between singing interest and dance interest. The first was represented in our final genetic model by paths from the residual C and E influences on instrument lessons to dance lessons, C path=.35, 95% CI [.05, .66], and E path=.11, 95% CI [.00, .28]. The second was represented by a path from the residual C influences on singing interest to dance interest, C path=.38, 95% CI [.20, .53]. These paths are displayed as correlations in Figure 2 for ease of viewing. Although we included these paths to prevent potentially biasing the genetic/environmental correlations between the latent factors upwards, results were nearly identical if they were not included in the model. The phenotypic residual correlations were also included in subsequent analyses with verbal ability.

Associations between Age 12 Music Engagement and Age 16 Verbal Ability

We next incorporated the multivariate model of musical instrument, singing, and dance engagement with age 16 verbal ability measures in a correlational phenotypic model. Latent variable correlations are displayed in Table 2, which also included correlations with age 12 IQ. Factor loadings on individual tasks/questionnaires are not displayed, but were nearly identical to previous analyses. This model fit the data well, $\chi^2(278)=501.32$, $p<.001$, RMSEA=.031, CFI=.970.

As shown in Table 2, the Music Instrument engagement factor was correlated with age 12 full-scale IQ, $r=.17$, 95% CI [.11, .23] for twins, $r=.21$, 95% CI [.14, .28] for siblings. Music Instrument engagement was also correlated with the age 16 Verbal Ability factor, $r=.23$, 95% CI [.17, .31]. The Singing or Dance factors were not associated with age 12 IQ or age 16 verbal ability. Similar associations were observed for verbal IQ and performance IQ, with the correlations between latent music engagement factors and verbal IQ nearly identical to those between music engagement and full-scale IQ.

In the phenotypic regression model (Figure 3), age 16 Verbal Ability remained associated with age 12 Musical Instrument engagement after controlling for age 12 IQ, $\beta=.09$, 95% CI [.04, .15] for twins, $\beta=.10$, 95% CI [.04, .16] for siblings. Similar results were obtained when we re-ran this regression model controlling for verbal IQ or performance IQ instead of full-scale IQ (supplement Figure S3). Additional post-hoc analyses confirmed that associations with Musical Instrument engagement were similar even when the Verbal Ability factor was split into separate Vocabulary ($r=.22$) and Verbal Fluency ($r=.20$) latent factors (supplement Table S5).

We next fit a multivariate genetic model to decompose the phenotypic associations between Music Instrument engagement and Verbal Ability into its genetic and environmental underpinnings, displayed in Figure 4a. The genetic correlation, which appeared to explain most of the phenotypic association, was nonsignificant, $r_g=.23$, 95% CI [−.02, .44], but both the genetic and shared environmental paths could not be removed without a significant drop in fit, $\chi^2(2)=17.28$, $p=.001$. Figure 4b displays the significant genetic correlation, $r_g=.31$, 95% CI [.14, .47], after removing the shared environmental correlation from the model (i.e., shared environmental influences did not explain any variance in Verbal Ability, and removing this correlation greatly improved power for the genetic correlation). Furthermore, heritabilities and nonshared environmental influences on music engagement ($a^2=.76$, $e^2=.12$) and verbal ability ($a^2=.88$, $e^2=.05$) were similar, and neither trait demonstrated significant shared environmental influences. These results suggest that music engagement and verbal ability have very similar modes of inheritance, which is the least optimal situation to test the direction of causation model. Therefore, although we estimated this model (see supplement Figure S4), we do not put strong emphasis on the results and focus our later discussion of causal implication instead on longitudinal models from the previous paragraph

Finally, Figure 5 displays the genetic and environmental decomposition of the cross-sectional association between instrument engagement and full-scale IQ at age 12. In this case, there was a significant genetic correlation between music instrument engagement and full-scale IQ, $r_g=.44$, 95% CI [.21, .77] in twins, $r_g=.80$, 95% CI [.35, 1.0] in siblings. Shared and nonshared environmental correlations were actually estimated in the opposite direction, and significant for nonshared environmental influences, $r_e=-.98$, 95% CI [−1.0, −.22] in twins, $r_e=-.59$, 95% CI [−.83, −.15] in siblings, $\chi^2(1)=6.50$, $p=.011$, suggesting most (if not all) of the phenotypic association between music instrument engagement and IQ at age 12 ($r=.17$) is explained by genetic influences.

Discussion

We employed a sizeable twin/adoptive sample design to conduct a novel genetic investigation of multiple modes of engaging with music during adolescence and to explore the developmental relationship between language and music, two uniquely human traits. Heritability estimates for engaging with musical instruments (78%), vocal music (i.e., singing 43%), and dance (66%) contribute to knowledge of the genetic structure of musicality traits. We also uncovered associations between adolescent self-reported music engagement and later verbal ability measures while controlling for earlier IQ, thus providing population-level evidence of a robust link between musicality and language traits and that converges with findings from smaller studies (Gordon et al., 2015c; Swaminathan & Schellenberg, 2019). Engaging with musical instruments at age 12 also predicted longitudinal growth of verbal ability at age 16, consistent with theoretical predictions about the influence of musicality on language development (Brandt et al., 2012; Ladanyi et al., 2020; Patel, 2014). Innovative use of an adoptive sibling design allowed us to further investigate environmental and genetic influences, strengthening our power to detect shared environmental influences and correlations (Plomin & DeFries, 1985) compared to the classic twin design (Martin, Eaves, Kearsey, & Davies, 1978).

Genetic Structure of Adolescent Self-Reported Music Engagement

Our initial results demonstrated that musical instrument engagement was highly heritable. The estimate of 78% is slightly higher than heritability estimates for other music traits (Butkovic, Ullen, & Mosing, 2015; Hambrick & Tucker-Drob, 2015; Seesjarvi et al., 2016). Our higher estimates may be due to the latent variable approach, which typically yields higher heritability estimates because measurement error is modeled as part of the estimate of nonshared environmental influences on individual measures. Heritability estimates were lower but significant for singing and dancing engagement factors. These genetic influences likely reflect the contribution of many hundreds or thousands of independent genetic effects (Chabris et al., 2015).

It was somewhat surprising that the latent factor for instrument engagement was only weakly associated with those for singing and dance engagement given the many neurobiological mechanisms shared between them (i.e., auditory-motor synchronization; Patel & Iversen, 2014). One possibility for the relatively weak associations between instrument engagement and singing/dancing engagement is that we used only a few measures with limited response options (including only a “yes/no” response for lessons). These data were collected before validated scales such as the Goldsmith’s Musical Sophistication Index (Mullensiefen et al., 2014) were developed. Nevertheless, the latent factors assessed here captured the shared variance among interest, lessons, and skill measures within each domain quite well, with strong heritability estimates on the latent factors. Moreover, singing and dancing correlated well with each other (r=.40), and demonstrated strong genetic overlap (r_g=.88), suggesting that measurement alone could not account for the weaker association between instrument engagement and the other domains. Furthermore, only musical instrument engagement, and not singing or dance engagement, was associated with verbal ability or IQ, providing further evidence that instrument engagement is distinct from other types of music engagement (see also Ireland, Iyer, & Penhune, 2019; Mansens, Deeg, & Comijs, 2018).

Previous work using the Goldsmith's Musical Sophistication Index in adults suggests that the "singing abilities" subtest is similarly correlated with music listening tests (melodic memory, beat perception) and personality traits (e.g., the big five) as other subscales ("active engagement", "perceptual abilities", "music training"), though correlations between singing and other subtests were of stronger magnitude than those observed here (Mullensiefen et al., 2014). Other work in a younger sample has shown that although singing interest loaded onto the same factor as instrument items, singing engagement items were captured by a different factor than instrument engagement (Coon & Carey, 1989). Thus, although instrument engagement may be quite distinct from singing and dance engagement in early adolescence, these traits may converge in adulthood (e.g., when individuals have had more time and opportunities to engage in additional types of musical endeavors beyond the ones they learned first).

In summary, these results complement prior work on the genetics of music aptitude, practice habits, and achievement (Butkovic et al., 2015; Coon & Carey, 1989; Hambrick & Tucker-Drob, 2015; Seesjarvi et al., 2016). They also demonstrate the validity of domain-type music engagement factors that encompasses interest, lessons and skill separately for each type of musical involvement (vocal, instrumental, dance).

Genetic Relationship Between Music Engagement and Language Abilities

This work also provides novel evidence for genetic associations between adolescent music engagement, IQ, and language. Individual differences in music and language traits are strongly associated (Gordon et al., 2015c; Woodruff Carr et al., 2014), and prior work on genetics of music and cognition have focused on non-verbal intelligence (Mosing et al., 2014b) rather than on specific verbal ability tests. In the current study, regression analyses suggested that musical instrument engagement was associated with verbal ability 4 years later, even when controlling for concurrent IQ. These findings confirm that music and language processes are deeply intertwined, and cannot be explained solely by covariance with previous verbal and nonverbal abilities. This conclusion converges with recent work showing that musical variables predict spoken language skills even after controlling for non-verbal IQ (Gordon et al., 2015c; Swaminathan & Schellenberg, 2019). The verbal measures in 16 year-olds consisted of vocabulary and verbal fluency, which reflect lexico-semantic aspects of language and are a reasonable proxy for broad language ability, but may also capture some separate variance from other levels of language processing/ability cited earlier (i.e., reading, grammar, second language acquisition). Vocabulary and verbal fluency measures were similarly related to the music engagement measures here.

This work is also relevant to the ongoing debate about whether associations between music and IQ are driven by shared predispositions (including genetic factors) or potential causal associations of music exposure enhancing IQ (Mosing et al., 2014b; Sala & Gobet, 2017). Cross-sectionally, we found that the positive association between music engagement and IQ was explained entirely through genetic rather than environmental influences. Environmental correlations were actually estimated in the opposite direction, though only the nonshared environmental correlation was significant. It is difficult to interpret this negative environmental correlation in light of a positive phenotypic correlation. Most importantly, the findings support the idea that essentially all of the positive phenotypic association between music engagement and IQ is explained by genetic influences, which explained a much larger proportion of the variance in instrument engagement and IQ (at least in twins). Therefore, although it would be interesting if the negative environmental correlation is replicated in future work, as it stands it is only predicting a small negative covariance that is overwhelmed by the larger genetic covariance.

This genetic association with concurrent IQ (and later verbal ability) is consistent with self-selection explanations where a common set of genetic influences predisposes some individuals towards both music engagement and strong IQ, with both constructs predicting later verbal ability. Other work has differentiated neural processes in fast versus slow learners of speech and music tasks (Zatorre, 2013), with implications for individual differences in predispositions playing a role in specific trajectories of language and music learning (Zatorre, 2013; Zuk & Gaab, 2018). That literature was rooted in earlier correlational work exploring neuroplasticity and transfer of music training to language task performance (Bidelman & Alain, 2015; Magne et al., 2006; Schön, Magne, & Besson, 2004). Recent approaches have re-evaluated such findings and moved away from arguments of transfer to those related to self-selection and gene-environment interplay (Schellenberg, 2015; Zatorre, 2013; Zuk & Gaab, 2018). Taken together, a possible interpretation of our findings is that adolescents drawn to engage in musical instrument training have genetic predispositions that also bolster acquisition of new linguistic material (i.e., vocabulary). These genetic influences are likely to be enriched for neural function in language-related networks, with acquisition mediated by neural plasticity. Moreover, these findings converge within the larger literature on music and language abilities, including well-known genetic correlations among various language processes (Hayiou-Thomas, 2008; Lee et al., 2018; Olson et al., 2013) that are also phenotypically associated with music ability and engagement phenotypes (Gordon et al., 2015c; Ozernov-Palchik et al., 2018; Piro & Ortiz, 2009). Thus, the genetic associations observed here may reflect potentially genetically driven associations between music and language traits more broadly (in line with Chevrud’s conjecture, which states that phenotypically correlated traits are likely to share genetic architecture; Sodini, Kemper, Wray, & Trzaskowski, 2018).

Beyond cross-sectional genetic associations with IQ, longitudinal evidence provides some support for weak causal associations, in which some additional variance in later verbal ability was captured by earlier music engagement even after controlling for the (genetic) association with IQ. Most of the correlation between age 12 music engagement and age 16 verbal ability (r=.23) was attenuated after controlling for age 12 IQ, but some unique variance was still explained by age 12 music engagement (β=.09 to .10). Results were nearly identical even when we controlled specifically for verbal IQ (supplement Figure S3), which includes similar subtests to those captured by the verbal ability factor 4 years later (i.e., vocabulary), suggesting music engagement may relate to language acquisition, a process that is continuing to unfold throughout adolescence. These findings converge with other studies showing that musicality predicts individual differences in multiple aspects of language (phonological processing; prosody; reading; grammar, and second language acquisition) that are foundational for academic and social success (Gordon et al., 2015c; Huss et al., 2011; Morrill, McAuley, Dilley, & Hambrick, 2015; Slevc & Miyake, 2006; Woodruff Carr et al., 2014).

These results from a relatively large sample of over 1600 individuals do not conclusively point to a causal association between instrument engagement and later language. Indeed, the reduced phenotypic association when controlling for IQ suggests much of this association is driven by self-selection (noted above). Moreover, there could be unmeasured variables associated with music engagement that could be driving the longitudinal association beyond IQ, including openness to experience and socioeconomic status (Corrigall & Schellenberg, 2015; Corrigall, Schellenberg, & Misura, 2013). The fact that instrument engagement but not singing or dance engagement (which may relate similarly to openness and require similar financial resources to enroll in lessons) was associated with later cognition potentially rules out these alternative explanations, but there could be other confounds specific to music instrument playing. Nevertheless, there is some evidence for direct associations beyond shared genetic liability which should be examined in even larger datasets, with specific attention to weaknesses in musicality as a risk factor for developmental speech and language disorders and a potential treatment strategy (Ladanyi et al., 2020; Schon & Tillmann, 2015).

Strengths and Limitations

The study combined measures of music engagement across twin and adoptive/biological sibling samples, greatly increasing power to detect shared environmental correlations compared to traditional twin methods (Plomin & DeFries, 1985). However, even with this large sample we did not have power to distinguish between additive and dominant genetic influences (Martin et al., 1978). Furthermore, we could not employ the direction of causation model (Figure S4) to more strongly shed light on correlated genetic/environmental influences versus direct causal associations because instrument engagement and verbal ability had similar genetic/environmental etiologies (Heath et al., 1993; Rasmussen, Ludeke, & Hjelmborg, 2019). We also did not examine gene-environment correlations or gene-by-environment interactions, which are relevant for music traits (Hambrick & Tucker-Drob, 2015; Wesseldijk, Ullen, & Mosing, 2019), and can bias heritability and genetic correlation estimates.

Biological sibling correlations in CAP were considerably smaller than DZ correlations in LTS, suggesting the simplifying assumption that environments are similar between twins and siblings may not hold in this case. However, these differences may be driven by the sex differences observed here (i.e., CAP includes mostly opposite sex siblings whereas LTS includes only same sex siblings). Indeed, after removing opposite sex siblings from CAP, biological sibling correlations for dance measures (which showed the strongest sex differences) were stronger than shown in Table 1 (*r*s=.22 to .50) and more similar to DZ estimates. Other differences may be explained by the fact that CAP siblings are not the same age and were a slightly earlier birth cohort.

The sample was primarily Caucasian Americans, and our findings may not generalize to other groups. However, the sample was representative of the Colorado population at the time of recruitment.

Finally, music engagement was assessed using latent factor models, increasing power and generalizability of the findings, but the individual indicators were based on single self-reported questionnaire responses. Although similar self-reported musicality variables are now commonly used in population-based research (Mullensiefen et al., 2014), future genetic studies could examine whether task-based measures of musicality (Law & Zentner, 2012; Seesjarvi et al., 2016; Ullén et al., 2014) are also associated with language outcomes. It will also be interesting to examine the stability of these genetic and environmental influences across childhood and adolescence, including their associations with language, to examine how music engagement maps onto different linguistic and cognitive trajectories.

Summary and Conclusions

In a longitudinal genetic study design in over 1600 participants, we quantified the heritability of adolescent self-report musical phenotypes and their genetic associations with verbal ability. Instrument engagement was highly heritable (a^2=.78) and was genetically correlated with verbal intelligence beyond shared general cognitive processes. Heritabilities for singing (a^2=.43) and dance engagement (a^2=.66) were more moderate and not related to verbal ability. The association between musical instrument engagement and later verbal ability aligns with targeted studies on musicality and language abilities in preschool and elementary-aged children (Ozernov-Palchik et al., 2018; Swaminathan & Schellenberg, 2019) and with recent population-based findings that training on an instrument in middle and high school predicts academic achievement (Guhn, Emerson, & Gouzouasis, 2020). We also found a robust longitudinal relationship between music engagement and language skills that suggests that music engagement in middle schoolers influences verbal task performance in high schoolers, in line with theoretical predictions about music engagement influencing language acquisition (and suggesting that both genetic pleiotropy and training are at play).

It will be important for future work to evaluate the extent to which these associations reflect shared neural architecture (and its underlying genetic influences), using both family-based genetic and genomic methods. These lines of research will help establish the dynamical relationship between musicality and speech-language traits throughout the lifespan (Mansens et al., 2018; Zhao & Kuhl, 2016) and in both typical and atypical populations (Ladanyi et al., 2020).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

This work was supported by supported by NIH grants R01AG046938, R01DC016977, DP2HD098859, K18DC017383, R01MH063207, R01DA046064, U01DA046413, and R01DA042742. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Bashir AS, & Scavuzzo A (1992). Children with language disorders: natural history and academic success. J Learn Disabil, 25, 53–65; discussion 66–70. doi:10.1177/002221949202500109

Bidelman GM, & Alain C (2015). Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. J Neurosci, 35, 1240–1249. doi:10.1523/JNEUROSCI.3292-14.2015

Bidelman GM, & Mankel K (2019). Reply to Schellenberg: Is there more to auditory plasticity than meets the ear? Proc Natl Acad Sci U S A, 116, 2785–2786. doi:10.1073/pnas.1900068116

Brandt A, Gebrian M, & Slevc LR (2012). Music and early language acquisition. Frontiers in Psychology, 3, 327. doi:10.3389/fpsyg.2012.00327

Brod G, & Opitz B (2012). Does it really matter? Separating the effects of musical training on syntax acquisition. Front Psychol, 3, 543. doi:10.3389/fpsyg.2012.00543

Butkovic A, Ullen F, & Mosing MA (2015). Personality related traits as predictors of music practice: Underlying environmental and genetic influences. Personality and Individual Differences, 74, 133–138. doi:10.1016/j.paid.2014.10.006

Chabris CF, Lee JJ, Cesarini D, Benjamin DJ, & Laibson DI (2015). The fourth law of behavior genetics. Current Directions in Psychological Science, 24, 304–312. doi:10.1177/0963721415580430

Clement S, Planchou C, Beland R, Motte J, & Samson S (2015). Singing abilities in children with Specific Language Impairment (SLI). Front Psychol, 6, 420. doi:10.3389/fpsyg.2015.00420

Coon H, & Carey G (1989). Genetic and environmental determinants of musical ability in twins. Behavior Genetics, 19, 183–193. doi:10.1007/bf01065903

Corley RP, Reynolds CA, Wadsworth SJ, Rhea SA, & Hewitt JK (2019). The Colorado Twin Registry: 2019 update. Twin Research and Human Genetics, 22, 707–715. doi:10.1017/thg.2019.50

Corrigall KA, & Schellenberg EG (2015). Predicting who takes music lessons: Parent and child characteristics. Frontiers in Psychology, 6, 282. doi:10.3389/fpsyg.2015.00282

Corrigall KA, Schellenberg EG, & Misura NM (2013). Music training, cognition, and personality. Frontiers in Psychology, 4, 222. doi:10.3389/fpsyg.2013.00222

de Diego-Balaguer R, Martinez-Alvarez A, & Pons F (2016). Temporal Attention as a Scaffold for Language Development. Front Psychol, 7, 44. doi:10.3389/fpsyg.2016.00044

Defries JC, Plomin R, Vandenberg SG, & Kuse AR (1981). Parent-offspring resemblance for cognitive-abilities in the Colorado Adoption Project: Biological, adoptive, and control parents and one-year-old children. Intelligence, 5, 245–277. doi:10.1016/S0160-2896(81)80012-8

Derks EM, Dolan CV, & Boomsma DI (2004). Effects of censoring on parameter estimates and power in genetic modeling. Twin Research, 7, 659–669. doi:10.1375/1369052042663832

Dittinger E, Barbaroux M, D'Imperio M, Jancke L, Elmer S, & Besson M (2016). Professional Music Training and Novel Word Learning: From Faster Semantic Encoding to Longer-lasting Word Representations. J Cogn Neurosci, 28, 1584–1602. doi:10.1162/jocn_a_00997

Fitch WT (2017). Empirical approaches to the study of language evolution. Psychonomic Bulletin & Review, 24, 3–33.

Gingras B, Honing H, Peretz I, Trainor LJ, & Fisher SE (2015). Defining the biological bases of individual differences in musicality. Philosophical Transactions of the Royal Society B: Biological Sciences, 370, 20140092. doi:10.1098/rstb.2014.0092

Gordon RL, Fehd HM, & McCandliss BD (2015a). Does music training enhance literacy skills? A meta-analysis. Frontiers in Psychology, 6, 1777. doi:10.3389/fpsyg.2015.01777

Gordon RL, Jacobs MS, Schuele CM, & McAuley JD (2015b). Perspectives on the rhythm-grammar link and its implications for typical and atypical language development. Annals of the New York Academy of Sciences, 1337, 16–25. doi:10.1111/nyas.12683

Gordon RL, Shivers CM, Wieland EA, Kotz SA, Yoder PJ, & Devin McAuley J (2015c). Musical rhythm discrimination explains individual differences in grammar skills in children. Developmental Science, 18, 635–644. doi:10.1111/desc.12230

Guhn M, Emerson SD, & Gouzouasis P (2020). A population-level analysis of associations between school music participation and academic achievement. Journal of Educational Psychology, 112, 308–328. doi:10.1037/edu0000431

Gustavson DE, Panizzon MS, Franz CE, Reynolds CA, Corley RP, Hewitt JK, … Friedman NP (2019). Integrating verbal fluency with executive functions: Evidence from twin studies in adolescence and middle age. Journal of Experimental Psychology: General, 148, 2104–2119. doi:10.1037/xge0000589

Hambrick DZ, & Tucker-Drob EM (2015). The genetics of music accomplishment: Evidence for gene-environment correlation and interaction. Psychonomic Bulletin & Review, 22, 112–120. doi:10.3758/s13423-014-0671-9

Hayiou-Thomas ME (2008). Genetic and environmental influences on early speech, language and literacy development. Journal of communication disorders, 41, 397–408. doi:10.1016/j.jcomdis.2008.03.002

Heath AC, Kessler RC, Neale MC, Hewitt JK, Eaves LJ, & Kendler KS (1993). Testing hypotheses about direction of causation using cross-sectional family data. Behavior Genetics, 23, 29–50. doi:10.1007/bf01067552

Honing H (2018). On the biological basis of musicality. Annals of the New York Academy of Sciences, 1423, 51–56. doi:10.1111/nyas.13638

Hu LT, & Bentler PM (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychological Methods, 3, 424–453. doi:10.1037//1082-989x.3.4.424

Huss M, Verney JP, Fosker T, Mead N, & Goswami U (2011). Music, rhythm, rise time perception and developmental dyslexia: perception of musical meter predicts reading and phonology. Cortex, 47, 674–689. doi:10.1016/j.cortex.2010.07.010

Hutchins S (2018). Early childhood music training and associated Improvements in music and language abilities. Music Perception: An Interdisciplinary Journal, 35, 579–593. doi:10.1525/mp.2018.35.5.579

Ireland K, Iyer TA, & Penhune VB (2019). Contributions of age of start, cognitive abilities and practice to musical task performance in childhood. PLoS One, 14, e0216119. doi:10.1371/journal.pone.0216119

Kaufman AS (1994). Intelligent testing with the WISC-III: John Wiley & Sons.

Kraus N, & White-Schwoch T (2016). Timescales of Auditory Processing. The Hearing Journal, 69, 36–40. doi:10.1097/01.HJ.0000479421.52441.9a

Ladanyi E, Persici V, Fiveash A, Tillmann B, & Gordon RL (2020). Is atypical rhythm a risk factor for developmental speech and language disorders? WIREs Cognitive Science, e1528. doi:10.1002/wcs.1528

Law LN, & Zentner M (2012). Assessing musical abilities objectively: construction and validation of the profile of music perception skills. PLoS One, 7, e52508. doi:10.1371/journal.pone.0052508

Lee T, Thalamuthu A, Henry JD, Trollor JN, Ames D, Wright MJ, … Team, O. R. (2018). Genetic and environmental influences on language ability in older adults: Findings from the Older Australian Twins Study. Behavior Genetics, 1–11. doi:10.1007/s10519-018-9897-z

Magne C, Schön D, & Besson M (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. Journal of Cognitive Neuroscience, 18, 199–211. doi:10.1162/jocn.2006.18.2.199

Mankel K, & Bidelman GM (2018). Inherent auditory skills rather than formal music training shape the neural encoding of speech. Proceedings of the National Academy of Sciences, 115, 13129–13134. doi:10.1073/pnas.1811793115

Mansens D, Deeg DJH, & Comijs HC (2018). The association between singing and/or playing a musical instrument and cognitive functions in older adults. Aging Ment Health, 22, 964–971. doi:10.1080/13607863.2017.1328481

Marie C, Magne C, & Besson M (2011). Musicians and the metric structure of words. J Cogn Neurosci, 23, 294–305. doi:10.1162/jocn.2010.21413

Martin NG, Eaves LJ, Kearsey MJ, & Davies P (1978). The power of the classical twin study. Heredity, 40, 97–116. doi:10.1038/hdy.1978.10

Mas-Herrero E, Marco-Pallares J, Lorenzo-Seva U, Zatorre RJ, & Rodriguez-Fornells A (2013). Individual differences in music reward experiences. Music Perception, 31, 118–138. doi:10.1525/Mp.2013.31.2.118

Morrill TH, McAuley JD, Dilley LC, & Hambrick DZ (2015). Individual differences in the perception of melodic contours and pitch-accent timing in speech: Support for domain-generality of pitch processing. Journal of Experimental Psychology: General, 144, 730–736. doi:10.1037/xge0000081

Mosing MA, Madison G, Pedersen NL, Kuja-Halkola R, & Ullen F (2014a). Practice does not make perfect: No causal effect of music practice on music ability. Psychological Science, 25, 1795–1803. doi:10.1177/0956797614541990

Mosing MA, Pedersen NL, Madison G, & Ullen F (2014b). Genetic pleiotropy explains associations between musical auditory discrimination and intelligence. PLoS One, 9, e113874. doi:10.1371/journal.pone.0113874

Mullensiefen D, Gingras B, Musil J, & Stewart L (2014). The musicality of non-musicians: an index for assessing musical sophistication in the general population. PLoS One, 9, e89642. doi:10.1371/journal.pone.0089642

Muthén LK, & Muthén BO (1998-2017). Mplus User's Guide: Eighth Edition. Los Angeles, CA: Muthén & Muthén.

Neale MC, & Cardon LR (1992). Methodology for genetic studies of twins and families. Dordrecht:: Kluwer Academic Publishers.

Neale MC, & Miller MB (1997). The use of likelihood-based confidence intervals in genetic models. Behavior Genetics, 27, 113–120. doi:10.1023/a:1025681223921

Olson RK, Hulslander J, Christopher M, Keenan JM, Wadsworth SJ, Willcutt EG, … DeFries JC (2013). Genetic and environmental influences on writing and their relations to language and reading. Annals of dyslexia, 63, 25–43. doi:10.1007/s11881-011-0055-z

Ozernov-Palchik O, Wolf M, & Patel AD (2018). Relationships between early literacy and nonlinguistic rhythmic processes in kindergarteners. Journal of Experimental Child Psychology, 167, 354–368. doi:10.1016/j.jecp.2017.11.009

Patel AD (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. Hear Res, 308, 98–108. doi:10.1016/j.heares.2013.08.011

Patel AD, & Iversen JR (2014). The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. Front Syst Neurosci, 8, 57. doi:10.3389/fnsys.2014.00057

Piro JM, & Ortiz C (2009). The effect of piano lessons on the vocabulary and verbal sequencing skills of primary grade students. Psychology of Music, 37, 325–347. doi:10.1177/0305735608097248

Plomin R, & DeFries JC (1983). The Colorado Adoption Project. Child Development, 54, 276–289.

Plomin R, & DeFries JC (1985). The origins of individual differences in infancy; the Colorado adoption project. Science, 230, 1369–1371.

Politimou N, Dalla Bella S, Farrugia N, & Franco F (2019). Born to speak and sing: Musical predictors of language development in pre-schoolers. Frontiers in Psychology, 10, 948. doi:10.3389/fpsyg.2019.00948

Rasmussen SHR, Ludeke S, & Hjelmborg JVB (2019). A major limitation of the direction of causation model: Non-shared environmental confounding. Twin Research and Human Genetics, 22, 14–26. doi:10.1017/thg.2018.67

Rebollo I, de Moor MHM, Dolan CV, & Boomsma DI (2006). Phenotypic factor analysis of family data: Correction of the bias due to dependency. Twin Research and Human Genetics, 9, 367–376. doi:DOI 10.1375/twin.9.3.367

Rhea SA, Bricker JB, Wadsworth SJ, & Corley RP (2013). The Colorado Adoption Project. Twin Research and Human Genetics, 16, 358–365. doi:10.1017/thg.2012.109

Rhea SA, Gross AA, Haberstick BC, & Corley RP (2013). Colorado Twin Registry: An update. Twin Research and Human Genetics, 16, 351–357. doi:10.1017/thg.2012.93

Sala G, & Gobet F (2017). When the music's over. Does music skill transfer to children's and young adolescents' cognitive and academic skills? A meta-analysis. Educational Research Review, 20, 55–67. doi:10.1016/j.edurev.2016.11.005

Sallat S, & Jentschke S (2015). Music Perception Influences Language Acquisition: Melodic and Rhythmic-Melodic Perception in Children with Specific Language Impairment. Behav Neurol, 2015, 606470. doi:10.1155/2015/606470

Satorra A, & Bentler PM (2001). A scaled difference chi-square test statistic for moment structure analysis. Psychometrika, 66, 507–514. doi:10.1007/Bf02296192

Schellenberg EG (2015). Music training and speech perception: A gene-environment interaction. Ann N Y Acad Sci, 1337, 170–177. doi:10.1111/nyas.12627

Schellenberg EG (2019). Music training, music aptitude, and speech perception. Proc Natl Acad Sci U S A, 116, 2783–2784. doi:10.1073/pnas.1821109116

Schön D, Magne C, & Besson M (2004). The music of speech: music training facilitates pitch processing in both music and language. Psychophysiology, 41, 341–349. doi:10.1111/1469-8986.00172.x

Schon D, & Tillmann B (2015). Short- and long-term rhythmic interventions: perspectives for language rehabilitation. Ann N Y Acad Sci, 1337, 32–39. doi:10.1111/nyas.12635

Seesjarvi E, Sarkamo T, Vuoksimaa E, Tervaniemi M, Peretz I, & Kaprio J (2016). The nature and nurture of melody: A twin study of musical pitch and rhythm perception. Behavior Genetics, 46, 506–515. doi:10.1007/s10519-015-9774-y

Slater J, Azem A, Nicol T, Swedenborg B, & Kraus N (2017). Variations on the theme of musical expertise: cognitive and sensory processing in percussionists, vocalists and non-musicians. European Journal of Neuroscience, 45, 952–963. doi:10.1111/ejn.13535

Slevc LR, & Miyake A (2006). Individual differences in second-language proficiency: does musical ability matter? Psychol Sci, 17, 675–681. doi:10.1111/j.1467-9280.2006.01765.x

Sodini SM, Kemper KE, Wray NR, & Trzaskowski M (2018). Comparison of genotypic and phenotypic correlations: Cheverud’s conjecture in humans. Genetics, 209, 941–948. doi:10.1534/genetics.117.300630

Soper DS (2018). A-priori sample size calculator for structural equation models [Software]. Available from http://www.danielsoper.com/statcalc.

Swaminathan S, & Schellenberg EG (2019). Musical ability, music training, and language ability in childhood. J Exp Psychol Learn Mem Cogn. doi:10.1037/xlm0000798

Swaminathan S, Schellenberg EG, & Venkatesan K (2018). Explaining the association between music training and reading in adults. J Exp Psychol Learn Mem Cogn, 44, 992–999. doi:10.1037/xlm0000493

Theorell T, Lennartsson AK, Madison G, Mosing M, & Ullén F (2015). Predictors of continued playing or singing–from childhood and adolescence to adult years. Acta Paediatrica, 104, 274–284.

Ullén F, Mosing MA, Holm L, Eriksson H, & Madison G (2014). Psychometric properties and heritability of a new online test for musicality, the Swedish Musical Discrimination Test. Personality and Individual Differences, 63, 87–93. doi:10.1016/j.paid.2014.01.057

Vinkhuyzen AA, van der Sluis S, Posthuma D, & Boomsma DI (2009). The heritability of aptitude and exceptional talent across different domains in adolescents and young adults. Behav Genet, 39, 380–392. doi:10.1007/s10519-009-9260-5

Wadsworth SJ, Corley RP, Munoz E, Trubenstein BP, Knaap E, DeFries JC, … Team, C. A. (2019). CATSLife: A study of lifespan behavioral development and cognitive functioning. Twin Research and Human Genetics, 1–12. doi:10.1017/thg.2019.49

Wechsler D (1974). Wechsler intelligence scale for Children - Revised. New York: Psychological Corporation.

Wechsler D (1981). Manual: Wechsler Adult Intelligence Scale - Revised. New York, N.Y.: Psychological Corporation.

Wechsler D (1997). Wechsler adult intelligence scale (3rd ed.). San Antinio, TX: Psychological Corporation.

Wesseldijk LW, Ullen F, & Mosing MA (2019). The effects of playing music on mental health outcomes. Scientific Reports, 9, 12606. doi:10.1038/s41598-019-49099-9

Woodruff Carr K, White-Schwoch T, Tierney AT, Strait DL, & Kraus N (2014). Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. Proc Natl Acad Sci U S A, 111, 14559–14564. doi:10.1073/pnas.1406219111

Zatorre RJ (2013). Predispositions and plasticity in music and speech learning: neural correlates and implications. Science, 342, 585–589. doi:10.1126/science.1238414

Zentner M, & Gingras B (2019). The assessment of musical ability and its determinants. Foundations of music psychology: Theory and research, 641–683.

Zhang JD, & Schubert E (2019). A single item measure for identifying musician and nonmusician categories based on measures of musical sophistication. Music Perception, 36, 457–467. doi:10.1525/mp.2019.36.5.457

Zhao TC, & Kuhl PK (2016). Musical intervention enhances infants' neural processing of temporal structure in music and speech. Proc Natl Acad Sci U S A, 113, 5212–5217. doi:10.1073/pnas.1603984113

Zuk J, & Gaab N (2018). Evaluating predisposition and training in shaping the musician's brain: the need for a developmental perspective. Annals of the New York Academy of Sciences. doi:10.1111/nyas.13737

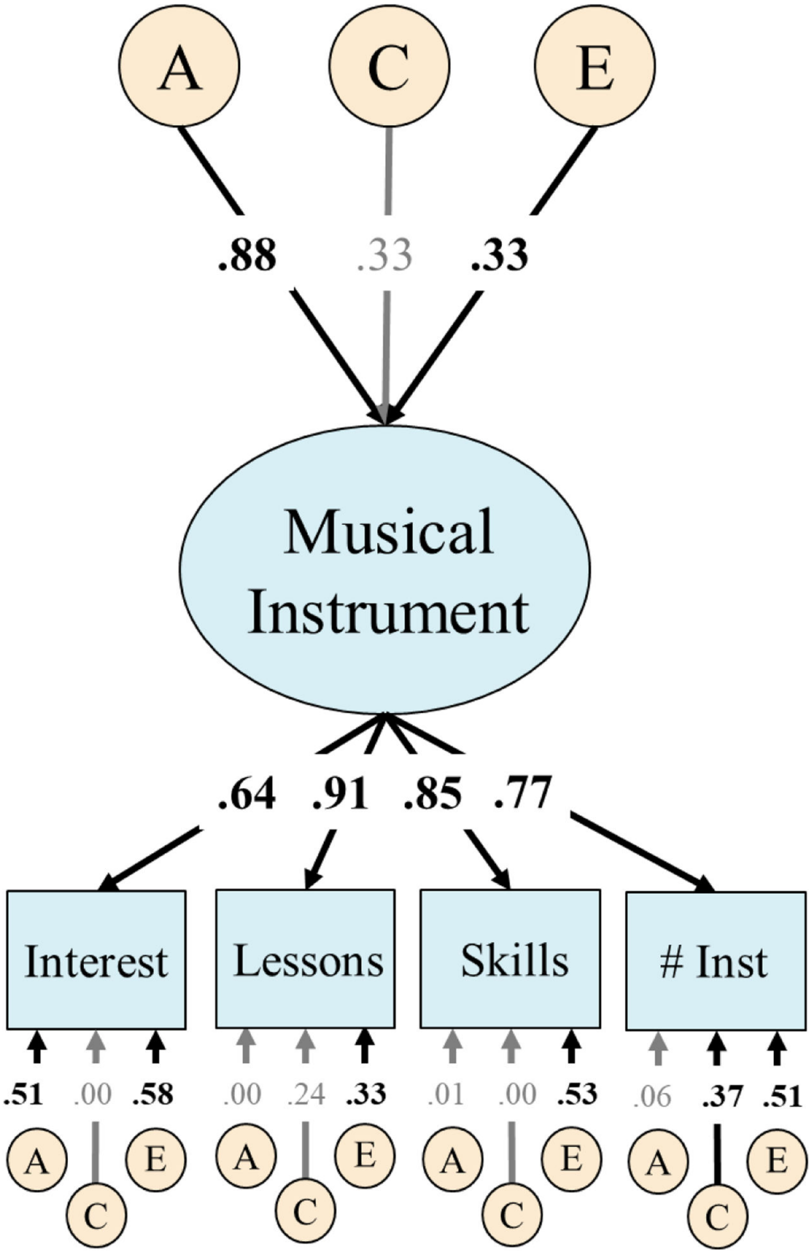

Figure 1:

Common pathway model of musical instrument engagement at age 12. The ACE factors represent the genetic (A), shared environmental (C), and nonshared environmental influences (E) on the latent variable (shown in an oval) or residual influences on the measured variables (in rectancles). Percent variance in explained of measured variables (boxes) or latent variables (ovals) can be computed by squaring factor loadings. This model fit the data well: $\chi^2(158) = 303.09$, p=.000, RMSEA=.068, CFI=.972. Significant factor loadings are displayed in bold and with black lines ($p < .05$). # Inst = Total number of instruments ever played (including singing).

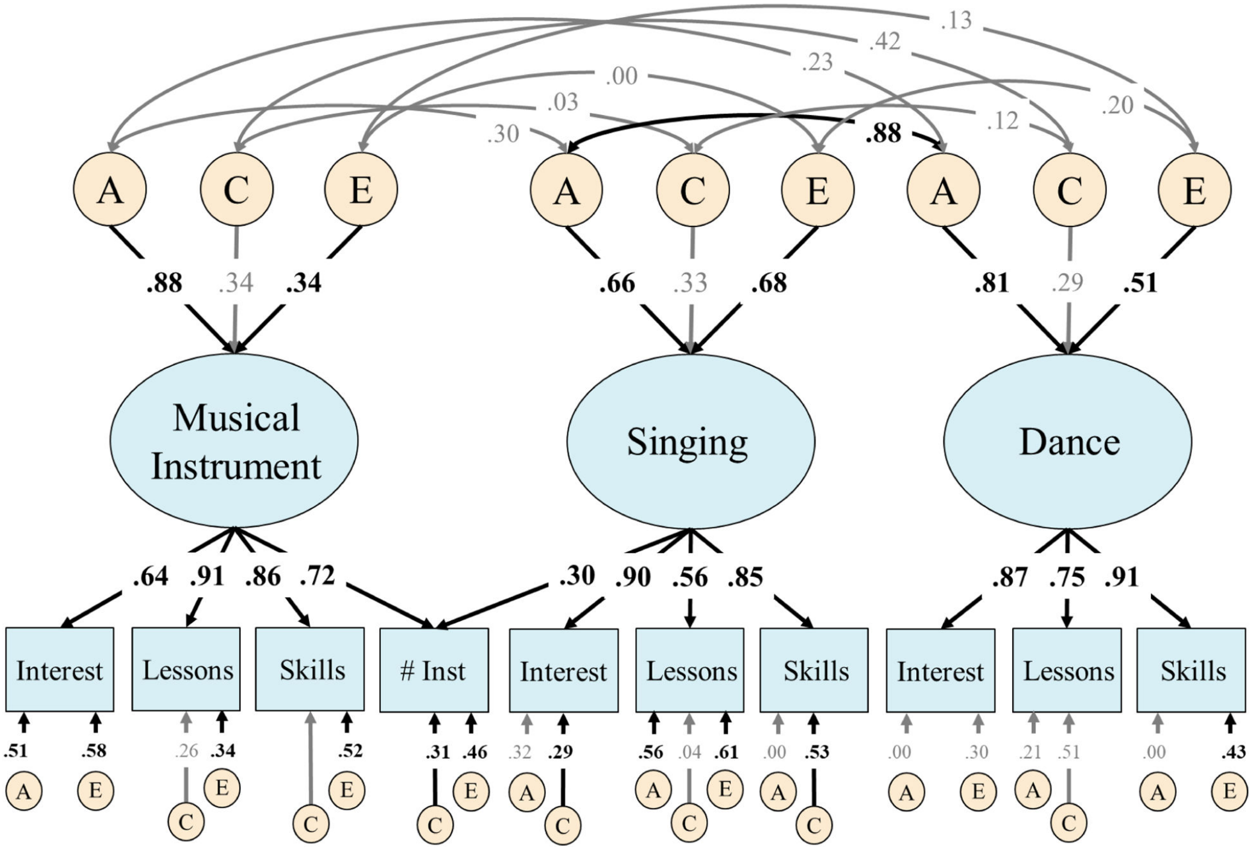

Figure 2:

Model of the genetic (A), shared environmental (C), and nonshared environmental influences (E) associations between music instrument, singing, and dancing engagement factors at age 12. Percent variance in explained of measured variables (boxes) or latent variables (ovals) can be computed by squaring factor loadings. Genetic and environmental correlations were not estimated directly but comptued from the Cholesky decomposition. This model fit the data well: $\chi^2(869) = 1179.36$, p=.000, RMSEA=.042, CFI=.967. Significant factor loadings are displayed in bold and with black lines ($p < .05$). # Inst = Total number of instruments ever played (including singing).

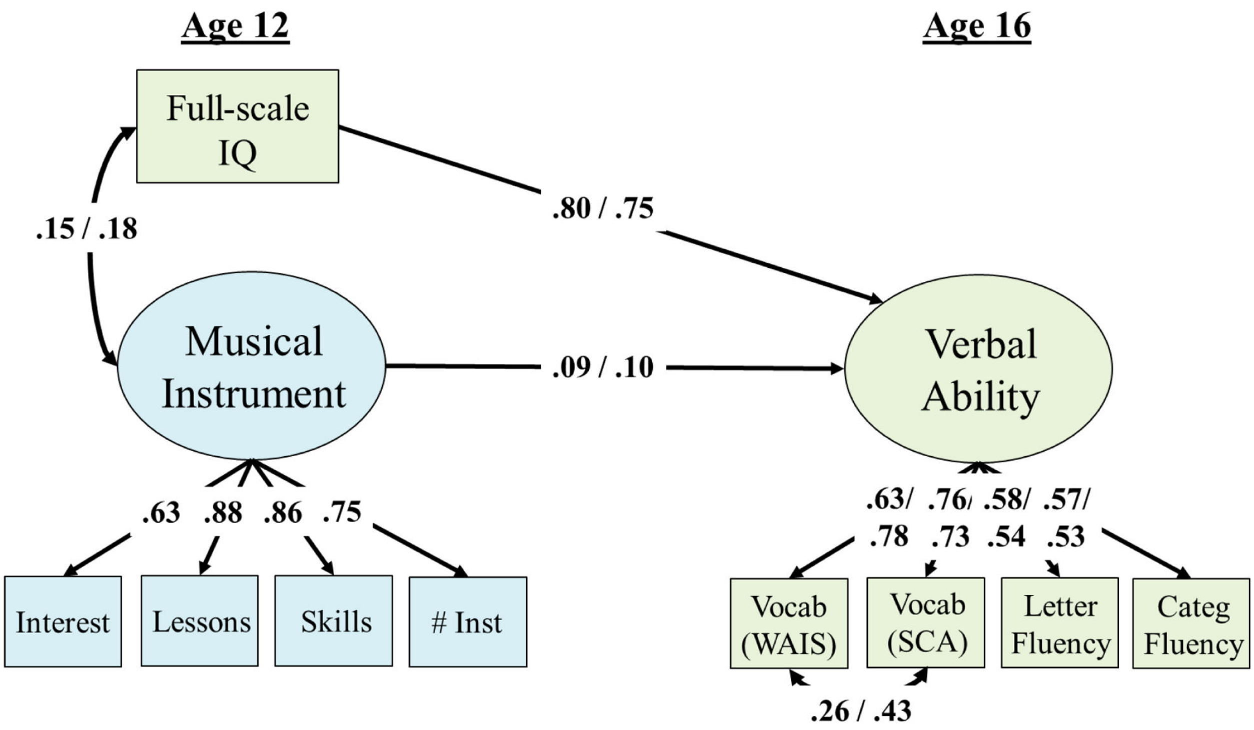

Figure 3:

Phenotypic regression models of musical instrument engagement at age 12 predicting verbal ability measures at age 16 controlling for age 12 full-scale IQ. The mean and residual variances for WAIS vocabulary are freed across the twin (displayed on the left) and adoptive subsamples (displayed on the right). The mean and residual variance for age 12 full-scale IQ is also freed across subsamples. Significant paths are displayed in bold, with black text and arrows (p<.05). Model fit: $\chi^2(73) = 203.25$, p=.000, RMSEA=.047, CFI=.967. # Inst = Total number of instruments ever played (including singing). WAIS = Wechsler Adult Intelligence Scale; SCA = Specific cognitive abilities battery.

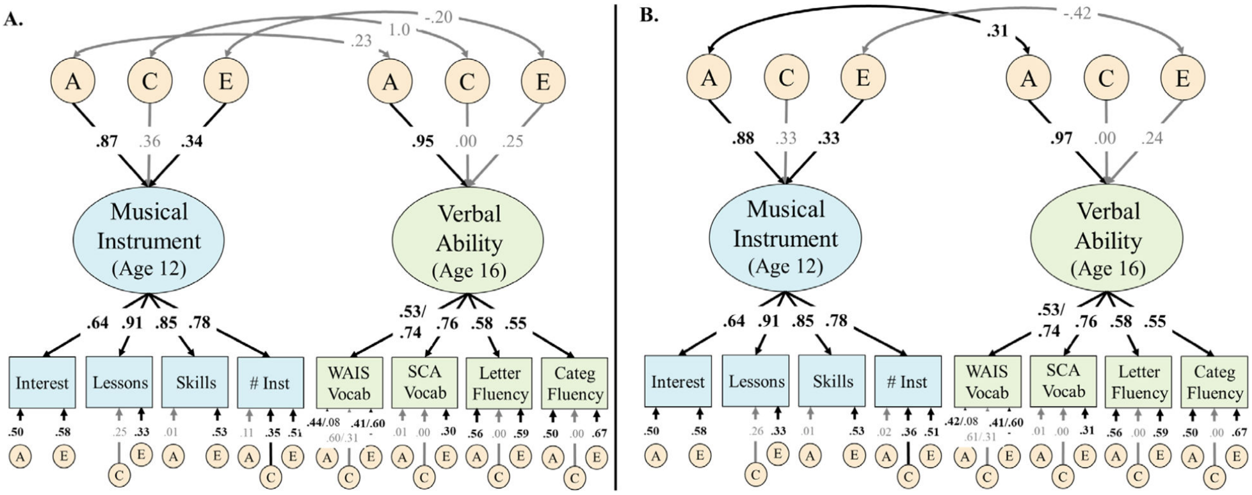

Figure 4:

Twin/adoptive model of musical instrument engagement at age 12 and verbal ability at age 16. Their association is modeled with genetic (A), shared environment (C), and nonshared environmental (E) correlations (Figure 4a) or only genetic and nonshared environmental correlations (Figure 4b; to improve power for the genetic correlation because shared environmental influences on verbal ability were estimated at 0). Figure 4b (with no C correlation) did not fit worse than Figure 4a, $\chi^2(1) = 0.99$, p=.320. In both models, the mean and residual variances for WAIS vocabulary were freed across subsamples, leading to different standardized estimates in each subsample (twins displayed on the left, adoptive/biological siblings displayed in the right). Not displayed are ACE residual paths from age 16 WAIS vocabulary to age 16 specific cognitive abilities (SCA) vocabulary that were included based on phenotypic analyses (.51 and .52 for A, .23 and .25 for C, .09 and .09 for E, respectively). Significant paths are displayed in bold, with black text and arrows ($p<.05$). Genetic and environmental correlations were not estimated directly but comptued from the Cholesky decomposition. Model fit for A: $\chi^2(561) = 773.39$, p=.000, RMSEA=.041, CFI=.957. B: $\chi^2(562) = 770.33$, p=.000, RMSEA=.041, CFI=.957. # Inst = Total number of instruments ever played (including singing). WAIS = Wechsler Adult Intelligence Scale; SCA = Specific cognitive abilities battery.

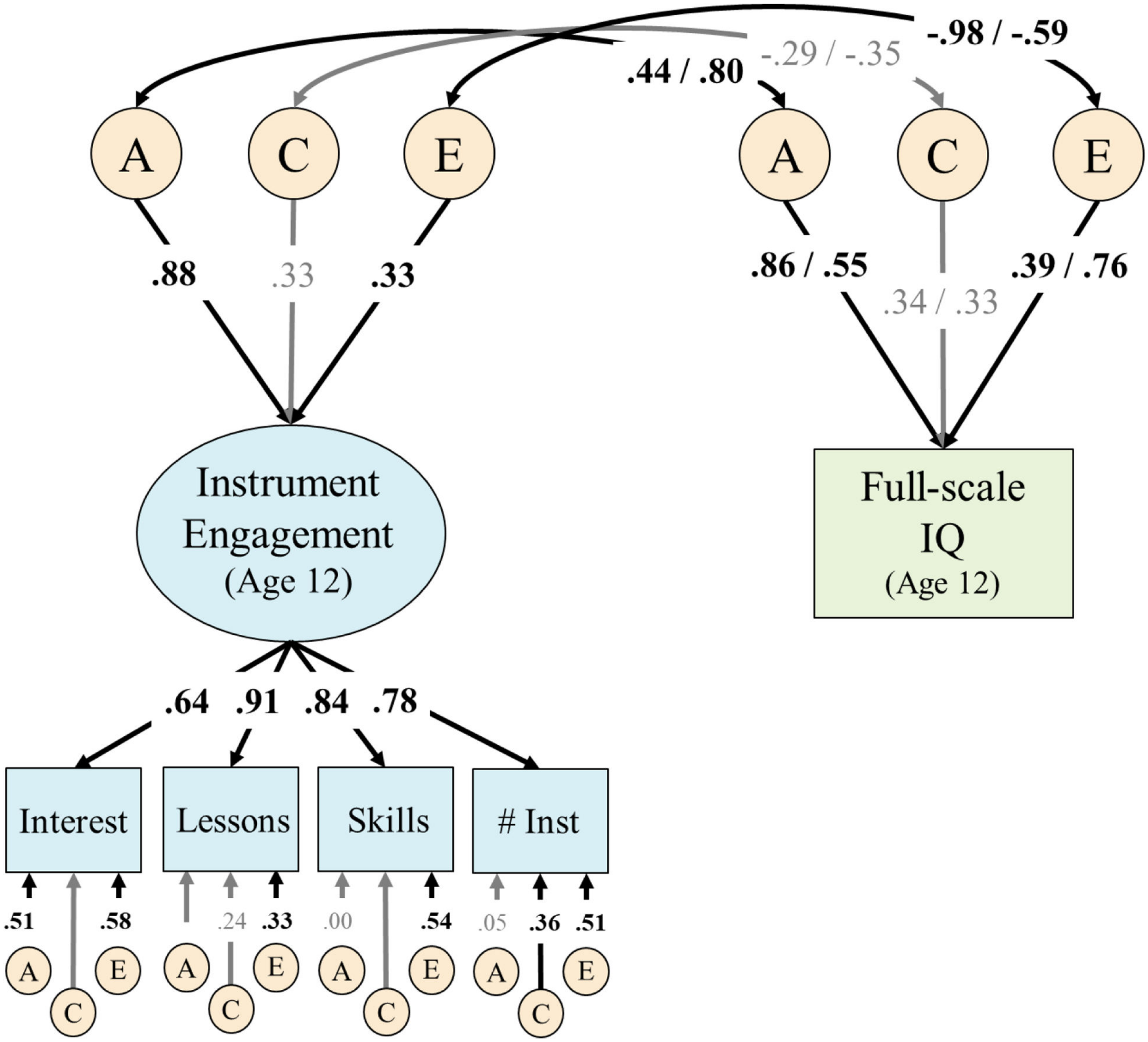

Figure 5:

Genetic and environmental correlations bewteen age 12 instrument engagement and age 12 full-scale IQ. The mean and residual variances for Verbal IQ was freed across the twin and adoptive subsamples, leading to different standardized estimates in each subsample (twins displayed on the left, adoptive/biological siblings displayed in the right). Significant paths are displayed in bold, with black text and arrows (*p*<.05). Genetic and environmental correlations were not estimated directly but comptued from the Cholesky decomposition. Model fit: $\chi^2(234) = 410.07$, p=.000, RMSEA=.061, CFI=.965. # Inst = Total number of instruments ever played (including singing).

Table 1:

Descriptive Statistics for All Study Measures

Task/Question	***N***	***M***	***SD***	**Range**	**Skewness**	**Kurtosis**	**rMZ**	**rDZ**
A. Longitudinal Twin Sample								
Age 12 Music Engagement								
Musical Instruments: Interest	642	2.31	0.74	1, 3	−0.57	−1.00	0.67	0.31
Musical Instruments: Lessons	721	0.62	0.49	0, 1	−0.48	−1.77	0.74	0.59
Musical Instruments: Skill	717	1.62	0.89	0, 3	−0.51	−0.52	0.61	0.52
Number of Instruments Played*	756	0.89	0.87	0, 6	1.33	3.64	0.63	0.53
Singing: Interest	663	1.84	0.79	1, 3	0.29	−1.33	0.50	0.44
Singing: Lessons	756	0.26	0.44	0, 1	1.10	−0.79	0.53	0.20
Singing: Skill	756	1.33	0.92	0, 3	0.07	−0.89	0.35	0.38
Dance: Interest	633	1.69	0.75	1, 3	0.59	−1.01	0.43	0.50
Dance: Lessons	757	0.29	0.45	0, 1	0.94	−1.12	0.84	0.81
Dance: Talent	757	1.12	0.93	0, 3	0.37	−0.81	0.47	0.48
Age 12 Intelligence (WISC-R or -III)								
Full-scale IQ	754	103.23	12.84	65, 136	0.00	−0.39	0.84	0.48
Verbal IQ	754	103.97	13.75	59, 145	−0.07	−0.10	0.83	0.55
Performance IQ	754	102.11	13.30	65, 142	0.09	−0.28	0.77	0.40
Age 16 Verbal Ability								
Vocabulary (WAIS-R or -III)	813	10.96	2.82	4, 19	0.06	−0.09	0.83	0.53
Vocabulary (SCA)	809	15.10	5.24	0.5, 36.5	0.61	0.72	0.86	0.61
Letter Fluency (SCA)	811	14.08	4.72	3.5, 29	0.45	0.06	0.63	0.34
Category Fluency (SCA)	811	7.05	2.74	0.5, 16.5	0.36	−0.05	0.61	0.33
B. Colorado Adoption Sample	*N*	*M*	*SD*	Range	Skewness	Kurtosis	rBIO	rADOP
Age 12 Music Engagement								
Musical Instruments: Interest	540	2.41	0.73	1, 3	−0.80	−0.72	0.06	−0.15
Musical Instruments: Lessons	583	0.68	0.47	0, 1	−0.80	−1.37	0.21	0.31
Musical Instruments: Skill	580	1.50	1.01	0, 3	−0.29	−1.08	0.20	−0.22
Number of Instruments Played*	610	1.06	0.89	0, 5	0.92	1.31	0.15	0.14
Singing: Interest	602	1.86	0.82	1, 3	0.27	−1.45	0.25	−0.06
Singing: Lessons	608	0.38	0.49	0, 1	0.49	−1.77	0.10	0.33
Singing: Skill	606	1.38	0.95	0, 3	−0.14	−1.04	0.13	−0.06
Dance: Interest	586	1.87	0.83	1, 3	0.25	−1.50	0.14	−0.06
Dance: Lessons	604	0.36	0.48	0, 1	0.57	−1.68	0.29	−0.11
Dance: Talent	601	1.27	1.04	0, 3	0.17	−1.20	0.01	−0.04
Age 12 Intelligence (WISC-R or -III)								
Full Scale IQ	613	111.18	11.32	72, 145	0.00	0.04	0.25	0.10
Verbal IQ	615	108.81	11.17	73, 145	0.12	0.34	0.32	0.03
Performance IQ	613	111.62	12.28	61, 150	0.00	0.08	0.27	0.18
Age 16 Verbal Ability								
Vocabulary (WAIS-R or -III)	776	8.59	2.06	4, 16	0.48	0.16	0.37	0.09

Task/Question	***N***	***M***	***SD***	**Range**	**Skewness**	**Kurtosis**	**rMZ**	**rDZ**
Vocabulary (SCA)	781	15.77	5.12	0, 34.5	0.54	0.71	0.28	0.06
Letter Fluency (SCA)	777	15.04	4.74	0, 30	0.25	−0.11	0.27	0.00
Category Fluency (SCA)	780	7.16	2.77	0, 17	0.33	−0.15	0.15	0.04

Note: The number of instruments played also includes singing. The final two columns display sibling correlations for monozygotic twins (*r*MZ), dizygotic twins (*r*DZ), biological siblings (*r*BIO) or unrelated siblings (*r*ADOP). WISC = Wechsler Intelligence Scale for Children; WAIS = Wechsler Adult Intelligence Scale; SCA = Specific cognitive abilities battery.

Table 2:

Latent Variables Correlations and 95% Confidence Intervals between Age 12 Music Engagement and Age 16 Verbal Ability

Latent Variables	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Age 12							
1. Instrument Engagement	1						
2. Singing Engagement	**.11**	1					
	[.03, .18]						
3. Dance Engagement	**.12**	**.40**	1				
	[.04, .19]	[.33, .47]					
4. Full-scale IQ	**.17 / .20**	.04 / .05	−.03 / −.03	1			
	[.11, .23] / [.13, .26]	[−.02, .10] / [−.02, .12]	[−.09, .03] / [−.10, .04]				
5. Verbal IQ	**.17 / .21**	.03 / .04	−.04 / −.05	**.69 / .97**	1		
	[.12, .23] / [.14, .28]	[−.03, .09] / [−.03, .11]	[−.10, .02] / [−.12, .02]	[.61, .77] / [.91, 1.0]			
6. Performance IQ	**.13**	.04	−.01	**.77 / .88**	**.43 / .52**	1	
	[.07, .19]	[−.02, .11]	[−.08, .05]	[.72, .82] / [.85, .91]	[.38, .47] / [.47, .58]		
Age 16							
7. Verbal Ability	**.23**	.04	−.07	**.70 / .80**	**.72 / .88**	**0.50**	1
	[.17, .31]	[−.04, .12]	[−.15, .01]	[.64, .75], [.75, .84]	[.66, .77] / [.83, .93]	[.44, .56]	

Note: Significant correlations are displayed in bold (based on 95% CIs). All constructs were measured with latent variables except verbal IQ and performance IQ at age 12. Means for age 12 verbal and performance IQ and age 16 WAIS vocabulary were freed across the LTS and CAP subsamples, as were the residual variances for age 12 verbal IQ and age 16 WAIS vocabulary, resulting in different standardized estimates in each subsample (twins displayed on the left, adoptive/biological siblings displayed in the right). Analyses controlled for effects of sex, ethnicity, race, and sibling type on each measure. Model fit: $\chi^2(307) = 556.31$, $p < .001$, RMSEA = .031, CFI = .967.