HHS Public Access
Author manuscript
J Child Psychol Psychiatry. Author manuscript; available in PMC 2021 June 01.

Published in final edited form as:
J Child Psychol Psychiatry. 2020 June ; 61(6): 689–698. doi:10.1111/jcpp.13134.

Early Classroom Reading Gains Moderate Shared Environmental Influences on Reading Comprehension in Adolescence

Jeanette Taylor[*],
Florida State University

Florina Erbeli[*],
Texas A&M University

Sara A. Hart,
Florida State University

Wendy Johnson
University of Edinburgh

Abstract

Background.—Reading is important for children's success in school and beyond yet many adolescents fail to reach expected levels of proficiency. This highlights the need to better understand the factors that influence reading effectiveness over time, including genes and environment. Greater expression of genetic influence on first and second grade reading fluency has been observed in higher quality classroom reading environments. To what degree this early environment continues to influence genetic and other environmental influences on later reading is unknown and was tested in this study.

Methods.—The quality of the early classroom reading environment was approximated by gains in oral reading fluency (ORF) across the school year among first- or second-grade classmates of 546 MZ and 1,016 DZ twin children (mean age = 7.13 years; *SD* = 0.45) who had reading comprehension scores from a state-wide mandatory test in school year 2013–14 when most twin pairs were in seventh to tenth grade (mean age = 14.41; *SD* = 1.13) in a variable called Class ORF Gain. Biometrical models were fit to the data to assess whether Class ORF Gain moderated the genetic, shared environmental, and/or non-shared environmental variance associated with adolescent reading comprehension.

Results.—Class ORF Gain moderated shared environmental influences on reading comprehension 6–9 years later. When early classroom reading gains were poor, variability in reading comprehension in adolescence was high and was associated largely with shared environmental influences. When early classroom reading gains were good, overall and shared environmentally influenced variability in adolescent reading comprehension was lower so that genetic influences were most relevant in explaining that variability.

Address correspondence about this manuscript to Jeanette Taylor, Department of Psychology, 1107 W. Call St., Florida State University, Tallahassee, FL, 32306-4301. ph. 850-644-7243; fax 850-644-7739; taylor@psy.fsu.edu.
[*]Co-first authors

Conflict of interest statement: no conflicts declared.

Conclusions.—Our findings suggested that classroom reading environment experienced when children were learning to read had a lasting influence on the factors underlying variability in later reading effectiveness.

Keywords

reading comprehension; genetic influence; environmental moderation

Attaining literacy is important for children's health and well-being. Children who master reading have better school outcomes that may, in turn, be associated with better social adjustment (Maughan, 1995), health (DeWalt, Berkman, Sheridan, Lohr, & Pignone, 2004), and future earnings as adults (McLaughlin, Speirs, & Shenassa, 2014). Conversely, children who fail to master reading are at risk for school dropout and associated negative consequences (Daniel et al., 2006). Reading skill does not emerge automatically – it requires active instruction followed by adequate practice to develop to expected levels of mastery. Formal instruction in reading begins in first grade in the U.S. However, children's reading proficiency does not progress uniformly, and around one-quarter of U.S. students in eighth and twelfth grades fall short of expected skill levels in critical areas such as comprehension (National Assessment of Education Progress, NAEP, 2017). This begs the question of *why*, with standard instruction in reading beginning in early school grades, some adolescents enter or even complete high school with deficient reading comprehension skills that could negatively impact their lives forever. The goal of the present study was to begin to address this question by examining the early learning environment's effect on adolescent reading comprehension.

Reading is learned through classroom instruction, but each child enters a learning environment as a unique individual who differs from her classmates on many factors, including prior learning experiences but also family genetic and environmental background. Genes and environment are important contributors to variation in reading skills in early learning stages (Christopher et al., 2013; Erbeli, Hart, & Taylor, 2018; Petrill, Deater-Deckard, Thompson, De Thorne, & Schatschneider, 2006), and genes appear to explain increasing proportions of variance in reading comprehension as children age (Little, Haughbrook, & Hart, 2017). Indeed, genetic differences account for half the variance in reading comprehension in adolescence (Tosto et al., 2017). However, finding evidence for genetic influence on a trait or behavior is not sufficient: it is important to understand the "interplay" or how genetic influences relate to environmental influences (Johnson, 2007).

The extent to which genes influence reading comprehension tends to differ with environment. For instance, one study observed that genetic influence on reading comprehension among third and fourth grade twins was moderated by school socioeconomic level, with children in poorer schools showing greater genetic influence on reading comprehension (Hart, Soden, Johnson, Schatschneider, & Taylor, 2013). Reading fluency (Taylor & Schatschneider, 2010) and general cognitive ability have been studied in children from age 7 (Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003) to late adolescence (Harden, Turkheimer, & Loehlin, 2007) and have shown the opposite pattern with regard to family-level socioeconomic status, such that heritability of those traits was greater in high-income homes. While these studies provide insight about gene-environment transactions underlying reading and related skills, they do not address environments directly related to the development of reading skills.

One key environment in the development of reading skills is the early classroom when children are first exposed to instruction on reading. In the U.S., first and second grade classrooms are characterized by their focus of instructional time on reading, the schools' selected reading curricula, and teacher skills in teaching children to read. Those same classrooms are also influenced by factors such as material resources (e.g., quality of textbooks; number and nature of supplemental reading books, quiet space for supplemental reading), the teacher's attitude toward reading instruction, and the particular mix of students in the class. The quality of the early classroom reading environment influences development of children's reading skills (National Reading Panel, NRP, 2000). The quality of instruction that children receive also varies (Connor et al., 2009) and accounts for around 5–10% of individual differences in reading (Connor, Morrison, Fishman, Schatschneider, & Underwood, 2007; NRP, 2000). As noted already, individual differences in reading comprehension are substantially genetically influenced (Little et al., 2017). However, despite the evidence that quality of instruction and classroom reading environment matter, individual differences are typically greater among children within classrooms than across classrooms. Thus, it is the interplay of classroom environment and genetic background that may offer a key to understanding why children differ in mastery of reading over time.

We previously observed a gene-environment interaction on reading fluency using data from monozygotic (MZ; identical) and dizygotic (DZ; fraternal) twin pairs (Taylor, Roehrig, Soden Hensler, Connor, & Schatschneider, 2010). Oral reading fluency (ORF) was measured in Florida classrooms three times throughout the school year beginning in first grade when explicit reading instruction started. What was referred to as teacher quality (Taylor et al., 2010) was characterized by Class ORF Gain (the gain in fluently read words per minute among classmates of twins during first or second grade). Importantly, this measure probably captured both quality of instruction and student characteristics (e.g., ability, socioeconomic status, motivation, etc.) but perhaps also selection effects (e.g., parent characteristics that influence which school their children attend), and were not completely independent of the twins' own (genetically influenced) characteristics as they too contributed to the overall class atmosphere, despite their own gain scores not being included in our measure. As such, we now view Class ORF Gain as an approximate indicator of the "early classroom reading environment" while recognizing that it is not a pure measure of environment and may show genetic influence as has been shown with other "environmental" measures (e.g., Kendler & Baker, 2007). Higher Class ORF Gains indicated greater average gain in reading in words per minute by the twins' classmates at the end of the school year after controlling performance at the beginning of the year; lower scores indicated less classroom-level gain (Taylor et al., 2010). Genetic influences on the twins' ORF scores at the end of first or second grade were greater when Class ORF Gain during that same year was high than when it was low (Taylor et al., 2010). Reading fluency is an early skill that recedes in importance relative to comprehension as children shift from learning to read to reading to learn. As such, we tested whether that early classroom reading "environment" had lasting effects on genetic and/or environmental influences on variability in reading comprehension in adolescence. Based on our prior observation, we expected to find moderation of genetic influences on reading comprehension such that heritability would be higher for twins who had been exposed to classrooms with greater Class ORF Gains.

Method

Participants

Details of the procedures for the Florida Twin Project on Reading are available elsewhere (Taylor, Hart, Mikolajewski, & Schatschneider, 2013; Taylor & Schatschneider, 2010). Briefly, twins were ascertained through a match on last name, birth date, and school grade from data in Florida's Progress Monitoring and Reporting Network (PMRN), an electronic repository for assessment data that was required for public schools under guidelines associated with the No Child Left Behind Act of 2002. Parents of children who matched on the aforementioned factors were contacted via mail to ask if their children were twins and, if so, to complete a 5-item zygosity questionnaire assessing physical similarity of the twins and provide consent to the use of their twins' reading achievement data from the PMRN. Participants in this study were a subset of those included in a prior study on Class ORF Gain that used data from school years 2004–05 through 2007–08 when the twins were in first or second grade (mean age of 7.13 years, *SD* = 0.45, N = 1,612 twins; Taylor et al., 2010). The present analysis included those twins who remained in Florida schools and may have completed the state-mandated Florida Comprehensive Assessment Test (FCAT) in reading in school year 2013–14, 6–9 years later when most twin pairs were in seventh to tenth grade. The 2013–14 school year was the last year Florida administered the FCAT, so it was the last year allowing complete consistency in measuring reading comprehension for this cohort. Fifty twins were not on the normal trajectory, because they were either a grade behind (N = 48) or skipped a grade (N = 2) relative to their co-twins in 2013–14. We excluded them to ensure that the outcome measure was equivalent for members of a twin pair (avoiding differences within twin pairs stemming from different grade-level exposure). This yielded a total of 1,562 twins for moderation analyses. Specifically, there were 759 complete twin pairs: 261 MZ pairs (138 female; 123 male) and 498 DZ pairs (125 same-sex female; 121 same-sex male; 252 opposite sex) as well as 44 individual twins (14 females; 30 males). The mean age was 14.45 (*SD* = 1.13). There was a significant difference in race/ethnicity between the excluded twins and the rest of the sample, $\chi^2(4, N = 1{,}586) = 34.58$, $p < .01$. Examination of sources for this difference via standardized residuals across all combinations of group and race/ethnicity showed that significantly more Black and significantly fewer White twins were not on the normal trajectory than expected by chance. This might mirror the disproportional underrepresentation of Black children receiving early intervention in special education (Morgan et al., 2015; Morgan, Farkas, Hillemeier, & Maczuga, 2012). The distribution of "not on the normal trajectory" status across the five race/ethnicity groups are presented in Figure SI1 in the online Supporting Information. The excluded twins also had significantly lower mean Class ORF Gain (*M* = 54.60, *SD* = 17.20) than those who were included (*M* = 67.72, *SD* = 24.89), $t(1610) = 3.70$, $p < .01$, d = 0.61. The sample used included 30 sixth graders (age *M* = 13.16, *SD* = 0.58), 455 seventh graders (age *M* = 13.23, *SD* = 0.43), 439 eighth graders (age *M* = 14.20, *SD* = 0.47), 387 ninth graders (age *M* = 15.15, *SD* = 0.47), and 251 tenth graders (age *M* = 16.15, *SD* = 0.47). The racial/ethnic composition of the sample was: 1.1% Asian, 25.1% Black, 33.3% Hispanic, 35.4% White, and the remainder were mixed or other race/ethnicity. This closely matches the race/ethnicity of students currently in Florida schools: 21.9% Black, 33.8% Hispanic, and 37.4% White.

The moderation analyses were conducted using full information maximum likelihood, which allows missing data. As such, the moderation models included 247 twins who had Class ORF Gain data but were missing data on FCAT in 2013–14, yielding a sample of 1,315 individuals with FCAT data in grades 6 through 10 (age M = 14.41; SD = 1.13). There were significant race/ethnicity differences between the group with adolescent FCAT and those without, χ^2 (4, N = 1,538) = 27.83, $p < .01$. Examination of standardized residuals across all combinations of group and race/ethnicity showed that significantly more mixed race/ethnicity twins were missing FCAT than expected by chance. Importantly, there were no significant differences for Class ORF Gain between the group with adolescent FCAT (M = 67.92, SD = 24.96) and those without (M = 66.65, SD = 24.55), t (1,560) = 0.74, p = .46, d = 0.05.

Measures

All reading tests in the PMRN were administered by trained teachers or school staff in state-wide standardized formats as part of each school's academic program. The PMRN included only total test scores. No raw (item-level) data for tests were available to the researchers.

DIBELS Oral Reading Fluency (ORF).—ORF is a measure of sight word reading and measures children's reading accuracy and speed when reading words in coherent text. Children read each of three grade-level passages aloud for 1 min. Words omitted, substituted, and hesitations of more than 3 s are scored as errors. Words self-corrected within 3 s are scored as accurate. The number of words read correctly in one min from each passage is recorded, and the median number correct from the three passages is taken as the final score. Overall alternate-form reliability is .95 in first and .98 in second grade (Good, Kaminski, Smith, & Bratten, 2001). ORF scores in first/second grade ranged from 0 to 215. It was administered three times during the school year (fall, mid-year, and spring) using different forms each time.

Florida Comprehensive Assessment Test in Reading (FCAT).—FCAT is a measure of reading comprehension. Students are asked to read narrative and expository passages and answer multiple-choice, short- and long-response items based on passage content. FCAT was administered once during a 10-day testing window in the spring of 2014. Cronbach's alphas were .89 (sixth grade), .90 (seventh and ninth grades), and .85 (eight and tenth grades), and the criterion-related validities with Stanford Achievement Test Series were .83, .83, .82, .79, and .80 in sixth through tenth grades, respectively (Florida Department of Education, 2007). Standard scores were used here and ranged from 167 to 302 (Florida Department of Education, 2014).

Data Analysis

Operationalization of early classroom reading environment.—We identified each twin's first or second grade classroom and fit regression models to estimate their classmates' score gains over the three ORF assessments and then averaged them (Taylor et al., 2010). This provided a measure of early classroom reading environment that did not rely on the twins' own scores. That is, the Class ORF Gain variable was calculated separately for each twin and was used as a twin-level variable that was based on growth in reading of their classmates. The reliability of the Class ORF Gain for this study was $r = .74$. Class ORF Gain was by no means a "pure" measure of environment and was probably subject to selection effects related to parent characteristics and was not completely independent of the twins' own (genetically influenced) characteristics, as they too contributed to the overall class atmosphere, despite their own gain scores not being included in our measure. Nonetheless, it did provide an indicator of the quality of the environment in which twins learned to read.

Correlations between observed variables and twin correlations.—We calculated the (Pearson) correlation between observed FCAT and Class ORF Gain for the entire sample for the subsamples of twins whose Class ORF Gain scores were at the normal distribution mean (level 0) and one and two *SDs* from the mean (−2, −1, 1, 2). Twin data can be used to decompose phenotypic variance into additive genetic (A), shared environmental (C), and non-shared environmental components as well as measurement error (E). Intraclass correlations for FCAT were calculated separately for MZ and DZ twins within levels of Class ORF Gain. These twin correlations yield preliminary information on the relative magnitudes of underlying influences of A, C, and E. Higher MZ than DZ correlations suggest genetic influence; similar MZ and DZ correlations suggest shared environmental influences; and MZ correlations less than 1 suggest non-shared environmental influences. Finally, we calculated cross-twin cross-trait correlations (here, FCAT of one twin correlated with Class ORF Gain of his/her co-twin) as an initial indicator of genetic and environmental influences on the correlation between FCAT and Class ORF Gain. Similar interpretations about sources of influence are made when examining cross-twin cross-trait correlations as for intraclass correlations.

Twin models examining gene-environment interaction and correlation.—Purcell (2002) introduced a model for testing moderation on the variance in A, C, and E while accounting for possible gene-environment correlation (the association of variability in the environment with variability in genetic influences). Johnson (2007) provided an overview of gene-environment interplay models, including the one used here. We estimated the model shown in Figure 1 using full information maximum likelihood in Mx (Neale, Xie, & Boker, 1998) based on data that were residualized for sex, age, and age-squared, using the analytic approach taken by, for example, Johnson, de Ruiter, Kyvik, Murray, & Sorensen, 2015. This model estimates the A, C, and E variance that is associated with the moderator (Class ORF Gain) as indicated by paths leading from the first set of A, C, and E factors to Class ORF Gain (e.g., a_{11}). Those same factors can also account for variance in the outcome (FCAT) that is in common with variance in Class ORF Gain as indicated by paths from the first set of factors to FCAT (e.g., a_{21}). Covariance of genetic and of environmental factors associated with both the moderator and outcome variable can be calculated by multiplying the path estimates stemming from one of the first set of factors. For instance, genetic covariation between Class ORF Gain and FCAT is derived by multiplying a_{11} and a_{21}. Covariation for C and E are comparably calculated. A second set of A, C, and E factors are modeled to account for unique variance in FCAT alone (e.g., a_{22}). Finally, the model also allows for moderation of A, C, and E variance that is common to Class ORF Gain and FCAT as denoted with the b_nM term on the covariance paths (e.g., $a_{21} + b_1M$). Moderation of A, C, and E variance that is unique to FCAT is likewise possible via a moderation term on the relevant path from the second set of A, C, and E factors to FCAT alone (e.g., $a_{22} + b_4M$). The interval between the measurement of Class ORF Gain and FCAT varied from 6–10 years; we assumed that whatever the effects of early classroom reading environment on comprehension levels were modeled (with all its underlying assumptions) as genetic and environmental variance that did not vary over the early adolescent period of observation encompassed in our data.

We first fit a full model with moderation on all sources of A, C, and E (b_1 – b_6 in Figure 1 allowed to vary) and then systematically compared it to reduced models. Reduced models were specified setting one or more variance parameters (a_{21}, a_{22}, c_{21}, etc.) and/or one or more moderation parameters (b_1, b_2, etc.) to zero. We accepted reduced models over the full model based on chi-squared difference tests, Akaike's Information Criterion (Akaike, 1987), and Sample-Size Adjusted BIC. AIC is considered particularly effective with twin samples of the size used in this report (Markon & Krueger, 2004). There is evidence that models with AIC values within two points of each other have similar levels of support as the best model (Burnham & Anderson, 2004), which suggests caution when rejecting models based on AIC in such circumstances. In those instances, a weighted AIC (W_i; Wagenmakers & Farrell, 2004) can be used to compare reduced models that have similar AIC values. The W_i value reflects the probability of the model being accurate given the particular data and models being tested. We reduced the models to the extents possible not to rule out possible indicated effects but to focus on those that appeared to be most important.

Gene-environment interplay can involve direct moderating influences on variance components of an outcome variable (implying gene-environment interaction in which environmental circumstances affect people with different genes differently), but it also can include correlations between the components of the two variables attributable to genetic and shared and non-shared environmental influences (denoted *r*A, *r*C, and *r*E), reflecting gene frequency that varies with environmental exposure, either passively through social stratification or actively through individual choice. These correlations offer important clues to the processes through which gene-environment interplay takes place. Genetic and environmental correlations are not estimated directly in the model (i.e., as a path estimate), but instead must be calculated from the estimated variance parameters for relevant paths. We used the best-fitting moderation model variance parameter estimates to calculate the reported genetic and environmental correlations.

Results

For descriptive purposes, Table 1 presents the means of reading comprehension assessed via the FCAT for MZ and DZ twins by level of Class ORF Gain (although the moderation models were fit to the entire sample and not by level of Class ORF Gain). Twins exposed to higher levels of Class ORF Gain did better, on average, in their reading comprehension as adolescents. Twins who were exposed to the poorest levels of Class ORF Gain had lower FCAT performance later on, although variability was high. The FCAT data were normally distributed across levels of Class ORF Gain, which provided confidence that any moderating effects were not artifacts of the data distributions. Again, for descriptive purposes, Table 2 shows that the intraclass correlations from MZ and DZ twins varied over levels of Class ORF Gain, suggesting that magnitudes of genetic and shared environmental influences might also vary with levels of Class ORF Gain. Shared environmental influences largely accounted for the phenotypic association between Class ORF Gain and FCAT ($r = .37$), as evidenced by similar cross-twin cross-trait correlations for MZ and DZ twins at most levels of Class ORF Gain (see Table 2).

Univariate estimates of heritability and environment have been reported previously for Class ORF Gain (Taylor et al., 2010) and FCAT (e.g., Erbeli et al., 2018). Consistent with those reports, the present data yielded the following estimates for Class ORF Gain: A = .16 (.95 CI = .06, .26), C = .63 (.95 CI = .52, .75), E = .21 (.95 CI = .18, .25) and for FCAT: A = .37 (.95 CI = .23, .50), C = .41 (.95 CI = .28, .55), and E = .22 (.95 CI = .18, .26).

Moderation model results for the best-fitting set of models are summarized in Table 3 (for results of all models tested see Table SI1 in Supporting Information). The full moderation model in which all unique and common A, C, and E sources of variance in FCAT were allowed to be moderated by variance in Class ORF Gain served as the base. As shown in Table 3, the model with no moderation fit significantly worse than it did, which meant that moderation was present. Similarly, the model with moderation on all unique A, C, and E paths but no common A, C, and E paths could be rejected, indicating that one or more common paths was moderated. The model does not handle this situation well, as it indicates that correlated genetic and environmental influences are confounded with main mean effects of the moderating variable. We tested all possible moderation models (see Table SI1) to do as much as possible to identify the nature of this confounding.

The best-fitting set of models included moderation on common C. As seen in Table 3, it was possible to reduce the number of unique moderating parameters without significant loss of model fit (as indicated by nonsignificant chi-square difference tests), but not to distinguish clearly between moderation on unique A and C on the basis of AIC given that the lowest AIC value was only 1 point less than the next lowest AIC value. The intraclass correlations in Table 2 suggested differences in genetic influence over levels of the moderator, providing support for the model with moderation on common C and unique A and C. However, the moderating parameter for unique A was not significant, $b_4 = .03$ (.95 CI = −.03, .11) though the unique C moderating parameter was significant in that model, $b_5 = -.11$ (.95 CI = −.19, −.03), and in the model with moderation on common C, $b_2 = -.07$ (.95 CI = −.12, −.02), and unique C, $b_5 = -.09$ (.95 CI = −.16, −.02). In addition to having the lowest AIC value, the W_i value was best for the common C and unique C moderation model and, therefore, we selected that model as best. We corroborated this selection by truncating the moderating variable to ensure that the extremes were not the major reasons for the observed effects. When the Class ORF Gain variable was trimmed, such that anyone below the 1st (n = 0) and above the 99th percentile (n = 51) was recoded to these boundaries, all the moderating effects were replicated, providing confidence in their robustness. Finally, linear and quadratic main effects models were tested to ensure that the data were showing moderation rather than quadratic main effects (see results in Table SI2 in Supporting Information). There were no quadratic main effects, providing further confidence in the moderation results. Figure 2 presents the unstandardized variance estimates from the selected best-fitting model: Class ORF Gain moderated the shared environmental influences on FCAT but also the shared environmental influences common to both Class ORF Gain and FCAT.

Consistent with our prior work (Taylor et al., 2010), we conducted the model-fitting analyses without regard to whether twins within a pair shared a classroom in first or second grade and, therefore, had the same score for Class ORF Gain. This is potentially important as genetic and environmental influences and main mean effects are completely confounded when twins have the same score on the moderating variable (by definition or by chance), and the field currently has no models capable of disentangling them. We reran the models including only twins that were in different classrooms in first or second grade as an additional check on our results. The large majority of MZ and DZ twin pairs, 77% and 83%, respectively, were in different first or second grade classrooms. Given the sample size, it is not surprising that this small difference in proportion was significant, χ^2 (1, N = 1,562) = 9.42, $p = .002$, and it could indicate systematic tendencies to place more similarly presenting twins together, whether by schools or parental request based on child preference or assessment of 'school readiness' though we observed no evidence that it did. The subsample of twins in different classrooms had decreased power for the modeling analyses to the point that clearly relevant paths, such as the one for genetic effects unique to FCAT, lost significance. Nonetheless, the model-fitting results were consistent with those from the full sample, as were the phenotypic and twin correlations (see a summary of best-fitting models in Table SI3 and correlations in Table SI4). The model with moderation on common C, $b_2 = -.08$ (.95 CI = −.14, −.03), and unique C, $b_5 = -.08$ (.95 CI = −.16, −.01), remained the best choice with the pattern of moderation the same as shown in Figure 2 using the full sample and the magnitudes of moderating effects were very similar. As an additional check on the results, we tested a two-group moderation model. The unconstrained model that allowed parameters to vary by whether twins shared a classroom, $-2LL = 5484.93$ ($df = 2832$), Sample Size Adjusted BIC = −2188.79, did not differ significantly in fit from the constrained model that equated parameters, $-2LL = 5487.66$ ($df = 2838$), Sample Size Adjusted BIC = −2197.87, $p = .84$, and fit better according to the Sample Size Adjusted BIC fit statistics.

Finally, the genetic and environmental correlations indicate the degrees to which these sources of variance mediated the association between Class ORF Gain and later reading comprehension. The best-fitting moderation model showed moderation of only shared environment, which meant that only the shared environmental correlation could vary with levels of Class ORF Gain (although correlational estimates may be highly similar or identical across all levels of the moderator even when there is moderation). The shared environmental correlations were large and almost invariant, ranging from .63 at the lowest level and .62 at the highest level of Class ORF Gain. The genetic correlation was .37 across levels of the moderator. The non-shared environmental correlation was zero, indicating independence of those influences on early classroom reading gains and adolescent reading comprehension.

Discussion

The life-long influence of attaining adequate reading skills on children's lives makes it critical to understand its development. Children enter the school system with particular shared environmental and genetic backgrounds as well as unique experiences that shape who they are and influence how they learn. The specific classroom reading environments they enter also range in quality as a function of the caliber of the reading curriculum, the skill of the teacher in implementing it, and even the other children in the classroom that year. A fundamental question is whether and how that matters and, if so, for how long? Prior work shows that quality of early instruction influences reading skills during the year in which children receive that instruction (NRP, 2000). The aim of this study was to estimate the long-lasting impact of early classroom reading gains as a proxy for environment on the genetic and environmental influences that contribute to individual differences in reading comprehension in adolescence. Our results suggested that, yes, early classroom reading gains did matter over time and impacted later reading comprehension skill that could, in turn, have lasting effects.

Previously, we observed that the quality of first or second grade classroom reading environment as measured by early gains in ORF by twins' classmates (Class ORF Gain) moderated genetic variance associated with the twins' own ORF scores for that same school year (Taylor et al., 2010). Specifically, genetic influence accounted for more of the variation in ORF scores for twins exposed to high quality classrooms. Here, we examined those same twins in adolescence and observed that the level of the first or second grade classroom reading gain was still positively and moderately associated ($r = .37$) with reading, but in a different reading skill: comprehension. That is, after 6–9 years, there was still a sizeable relationship between the level of classroom reading gains experienced in first or second grade and how well adolescents understood what they read. On the surface, this suggested that Class ORF Gain was an indicator of the quality of the early classroom reading environment with lasting impact on reading comprehension, but shared environmental mediation suggested involvement of school-, neighborhood-, and family-level influences beyond the classroom.

The longitudinal gene-environment moderation models suggested a complex level of gene-environment interplay. Learning to read in a poor early classroom reading environment was associated with wider variability in reading comprehension several years later, with pertinent influences dominated by shared environment in a ratio of almost 3 to 1 over genetic influence. Contrary to expectations that were based on our prior study (Taylor et al., 2010), genetic influences remained stable across levels of Class ORF Gain. The twin correlations suggested that genetic influences were different across levels of the moderator, but the balance of the evidence from the model-fitting analyses favored the model without moderation of genetic influences. Interestingly, when the early classroom reading gains were good, the variability in later reading comprehension was constricted and shared environmental variance dropped to about half the magnitude of genetic variance. Thus, genetic influences were not moderated, but they *were* the largest source of variance in reading comprehension among twins who had experienced the highest levels of early classroom reading gains. This cross-over interaction indicated a dramatic difference in the relative influences of genes and shared environment associated with reading comprehension in later school grades based on the level of the early classroom reading gains when children were first learning to read. Notably, when using only data from twins who were in different classrooms in first or second grade, the results were very similar to those from the full sample. This suggests that some of the effects involve the broader early school environment including the school and surrounding neighborhood.

To interpret the gene-environment interplay fully, the observed interaction must be considered in conjunction with the genetic and environmental correlations. Shared environment was the strongest mediator, likely reflecting that students cluster in schools and surrounding neighborhoods based partly on family-wide reading potential (genetic background, parental socioeconomic status, and parenting environment). This contributes to an association between overall student potential and school quality that also creates a broad social environment that affects all students. The genetic correlation between Class ORF Gain and FCAT was constant across levels of Class ORF Gain and was similar in magnitude to the overall phenotypic correlation between FCAT and Class ORF Gain. This genetic correlation had three possible sources, all of which are likely and undermine interpretation of Class ORF Gain as solely a measure of quality of early classroom instruction. First, the genetic correlation likely indicated that children clustered in schools based partly on genes influencing reading development, creating population-level stratification of reading-related genes by school quality (consistent with the considerable stratification of schools by socioeconomic status found in the U.S.). Second, response to early reading instruction was likely linked with relevant genetic background, and, over time, students who had responded well took increasing initiative in developing their own reading comprehension skills (i.e., niche-picking; Johnson, 2007). Third, the “tone” set by early reading instruction quality and classroom performance levels likely influenced children’s individual reading aspirations, thereby affecting the effort they made to develop reading skills. The genetic and shared environmental correlations should not, however, be considered distinct. Brighter and higher socioeconomic status parents tend to impart their own values of educational attainment to their children along with their genes. They read to them more, talk to them in more intellectually stimulating ways, engage them in more intellectually stimulating activities, etc., creating family-level genetic and shared environmental correlations that tend to be reflected as genetic influences in models such as those we applied.

Exposure to low levels of early classroom reading gains was associated with lower mean performance in reading comprehension in adolescence, but the variability was greater, suggesting that not all children in poor early environments were on pathways to lower achievement. Poor early classroom reading gains appeared to increase the impact of a multitude of shared environmental influences on reading comprehension in adolescence. Some factors (e.g., family SES) may be less subject to intervention than others (e.g., parent attitudes toward education). Lower-performing schools could be targets for interventions to help set up supportive environments at home. This might include parent training workshops that teach parents how to manage time for homework and emphasize the value of education to increase parent aspirations for their children's education. Exposure to greater early classroom reading gains was associated with higher mean performance in reading comprehension in adolescence, and the variability was lower, suggesting that many of those children were launched early on a pathway to higher achievement. Thus, the observed variation in magnitude of shared environmental influence suggests that interventions at the home or school levels might have the most potential to help those who need them most.

The present results should be considered in light of some limitations. First, the nature of the indicated moderating effects suggested that correlated genetic and environmental influences were confounded with main mean effects in a manner that available models cannot disentangle. We did as much as possible to identify the nature of this confounding, but future work is needed to clarify it further. Second, as noted previously, the model with moderation on common C and unique A and C was similar in fit to the best-fitting model that did not include moderation of unique A. The sample size was powered to detect moderation, but not necessarily to pick up small moderation of genetic effects in the presence of significant shared environment moderation and clear evidence of confounded correlations between genetic and environmental influences and main mean effects. Third, although the model-fitting procedure utilizes the full range of the data in estimating parameters, there were very few twins in the lowest quality early classroom reading environments. The low numbers gave rise to the likelihood of restriction of range for the correlations calculated at the low end of Class ORF Gain. A final limitation is that early classroom reading environment was characterized by class performance on reading fluency. This is an important reading skill in first and second grade, but class performance on this skill is only one aspect of the quality of the classroom reading environment. Future research could improve upon the present work by measuring early classroom reading environment more comprehensively (e.g., considering teacher qualifications in reading instruction). Future research could also examine the influence of early classroom reading environment on other outcomes like overall academic achievement and math performance to assess whether effects generalize beyond reading.

Conclusions

The early classroom environment is a critical launching point for reading education, but it is not solely responsible for a child's success or failure. Children who experience poor classroom environment when learning to read are not all doomed to lower reading comprehension as adolescents. Their family environments and genetic backgrounds also clearly matter. Still, the advantages of good early classroom reading environments appear to pave the way for children to develop their reading ability fully and excel as readers later on – perhaps even in spite of suboptimal family environments along the way. Understanding why early school instruction on reading does not translate into adequate reading comprehension in later grades for all children requires an appreciation of intertwining genetic and environmental influences over time, potential shifts in the balance of impacts of genes and shared environment on developing reading skill, and the uniqueness of every student's pathway.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Jeanette Taylor and Sara A. Hart, Department of Psychology, Florida State University.

Sara A. Hart, Florida Center for Reading Research, Florida State University.

Florina Erbeli, Department of Educational Psychology, Texas A&M University.

Wendy Johnson, Department of Psychology, University of Edinburgh, UK.

This research was supported in part by grant NICHD P50 HD052120.

References

Akaike H (1987). Factor analysis and AIC. Psychometrika, 52, 317–332.

Burnham KP & Anderson DR (2004). Multimodel inference understanding AIC and BIC in model selection. Sociological Methods & Research, 33, 261–304.

Christopher ME, Hulslander J, Byrne B, Samuelsson S, Keenan JM, Pennington B, DeFries JC, Wadsworth SJ, Willcutt EG, & Olson RK (2013). Modeling the etiology of individual differences in early reading development: Evidence for strong genetic influences. Scientific Studies of Reading, 17, 350–368. [PubMed: 24489459]

Connor CM, Morrison FJ, Fishman BJ, Schatschneider C & Underwood P (2007). Algorithm-guided individualized reading instruction. Science, 315, 464–465. [PubMed: 17255498]

Connor CM, Piasta SB, Fishman B, Glasney S, Schatschneider C, Crowe E, Underwood P, & Morrison FJ (2009). Individualizing student instruction precisely: Effects of child by instruction interactions on first graders' literacy development. Child Development, 80, 77–100. [PubMed: 19236394]

Daniel SS, Walsh AK, Goldston DB, Arnold EM, Reboussin BA, & Wood FB (2006). Suicidality, school dropout, and reading problems among adolescents. Journal of Learning Disabilities, 39, 507–514. [PubMed: 17165618]

DeWalt DA, Berkman ND, Sheridan S, Lohr KN, & Pignone MP (2004). Literacy and health outcomes: a systematic review of the literature. Journal of General Internal Medicine, 19, 1228–1239. [PubMed: 15610334]

Erbeli F, Hart SA, & Taylor J (2018). Longitudinal associations among reading related skills and reading comprehension: A twin study. Child Development, 89, e480–e493. doi:10.1111/cdev.12853. [PubMed: 28569388]

Florida Department of Education. (2007). Assessment and accountability briefing book. (Tallahassee, FL: Author) http://www.fldoe.org/core/fileparse.php/7478/urlt/briefingbook07web.pdf).

Florida Department of Education (2014). 2014 FCAT 2.0. Tallahassee, FL: Author http://www.fldoe.org/core/fileparse.php/5668/urlt/0066933-2014fcat20_media.pdf).

Good RH, Kaminski R, Smith MR & Bratten J (2001). Technical adequacy and second grade DIBELS Oral Reading Fluency (DORF) passages (Tech. Rep. No. 8, Eugene: Univ. Oregon).

Harden KP, Turkheimer E, & Loehlin JC (2007). Genotype by environment interaction in adolescents' cognitive aptitude. Behavior Genetics, 37, 273–283. [PubMed: 16977503]

Hart SA, Soden B, Johnson W, Schatschneider C, & Taylor J (2013). Expanding the environment: gene × school-level SES interaction on reading comprehension. Journal of Child Psychology and Psychiatry, 54, 1047–1055. [PubMed: 23725549]

Johnson W (2007). Genetic and environmental influences on behavior: capturing all the interplay. Psychological Review, 114, 423–440. [PubMed: 17500633]

Johnson W, de Ruiter I, Kyvik KO, Murray AL & Sorensen TIA (2015). Genetic and environmental transactions underlying the associations among physical fitness/physical exercise and body composition. Behavior Genetics, 45, 84–105. [PubMed: 25359086]

Kendler KS & Baker JH (2007). Genetic influences on measures of the environment: A systematic review. Psychological Medicine, 37, 615–626. [PubMed: 17176502]

Little CW, Haughbrook R & Hart SA (2017). Cross-study differences in the etiology of reading comprehension: A meta-analytical review of twin studies. Behavior Genetics, 47, 52–76. [PubMed: 27630039]

Markon KE & Krueger RF (2004). An empirical comparison of information-theoretic selection criteria for multivariate behavior genetic models. Behavior Genetics, 34, 593–610. [PubMed: 15520516]

Maughan B (1995). Annotation: long-term outcomes of developmental reading problems. Journal of Child Psychology and Psychiatry, 36, 357–371. [PubMed: 7782402]

McLaughlin MJ, Speirs KE, & Shenassa ED (2014). Reading disability and adult attained education and income: evidence from a 30-year longitudinal study of a population-based sample. Journal of Learning Disabilities, 47, 375–386.

Morgan PL, Farkas G, Hillemeier MM, & Maczuga S (2012). Are minority children disproportionately represented in early intervention and early childhood special education? Educational Researcher, 41, 339–351. [PubMed: 24683265]

Morgan PL, Farkas G, Hillemeier MM, Mattison R, Maczuga S, Li H, & Cook M (2015). Minorities are disproportionately underrepresented in special education: Longitudinal evidence across five disability conditions. Educational Researcher, 44, 278–292. [PubMed: 27445414]

National Assessment of Educational Progress. (2017). The Condition of Education 2017. (NAEP; https://nces.ed.gov/programs/coe/pdf/coe_cnb.pdf).

National Reading Panel (2000). Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. (NIH Pub. 00–4769).

Neale MC, Xie G & Boker SM (1998). Introduction to the Mx Graphical User Interface Richmond, VA: Virginia Commonwealth Univ.

Petrill SA, Deater-Deckard K, Thompson LA, De Thorne LS & Schatschneider C (2006). Reading skills in early readers: genetic and shared environmental influences. Journal of Learning Disabilities, 39, 48–55. [PubMed: 16512082]

Purcell S (2002). Variance components models for gene-environment interaction in twin analysis. Twin Research, 5, 554–571. [PubMed: 12573187]

Taylor JE, Hart SA, Mikolajewski AJ & Schatschneider C (2013). An update on the Florida State Twin Registry. Twin Research & Human Genetics, 16, 471–475. [PubMed: 23067863]

Taylor J, Roehrig AD, Soden Hensler B, Connor CM & Schatschneider C (2010). Teacher quality moderates the genetic effects on early reading. Science, 328, 512–514. [PubMed: 20413504]

Taylor J & Schatschneider C (2010). Genetic influence on literacy constructs in kindergarten and first grade: Evidence from a diverse twin sample. Behavior Genetics, 40, 591–602. [PubMed: 20563747]

Tosto MG, Hayiou-Thomas ME, Harlaar N, Prom-Wormley E, Dale PS, & Plomin R (2017). The genetic architecture of oral language, reading fluency, and reading comprehension: a twin study from 7 to 16 years. Developmental Psychology, 53, 1115–1129. [PubMed: 28541066]

Turkheimer E, Haley A, Waldron M, D'Onofrio B, & Gottesman II (2003). Socioeconomic status modifies heritability of IQ in young children. Psychological Science, 14, 623–626. [PubMed: 14629696]

Wagenmakers E, & Farrell S (2004). AIC model selection using Akaike weights. Psychonomic Bulletin & Review, 11, 192–196. [PubMed: 15117008]

Key Points:

- Genes influence reading effectiveness beginning in childhood, but reading skills only develop with the help of an environmental factor: explicit instruction.
- Shared environmental influences explained most of the variation in adolescent reading comprehension but only for students exposed to poor classroom reading environments in first or second grade. Genetic influences explained more variability in adolescent reading comprehension for students exposed to good early classroom reading environments.
- The shared environment (family, school, neighborhood, etc.) may be an important area to focus intervention if students are to attain adequate reading comprehension levels as adolescents.

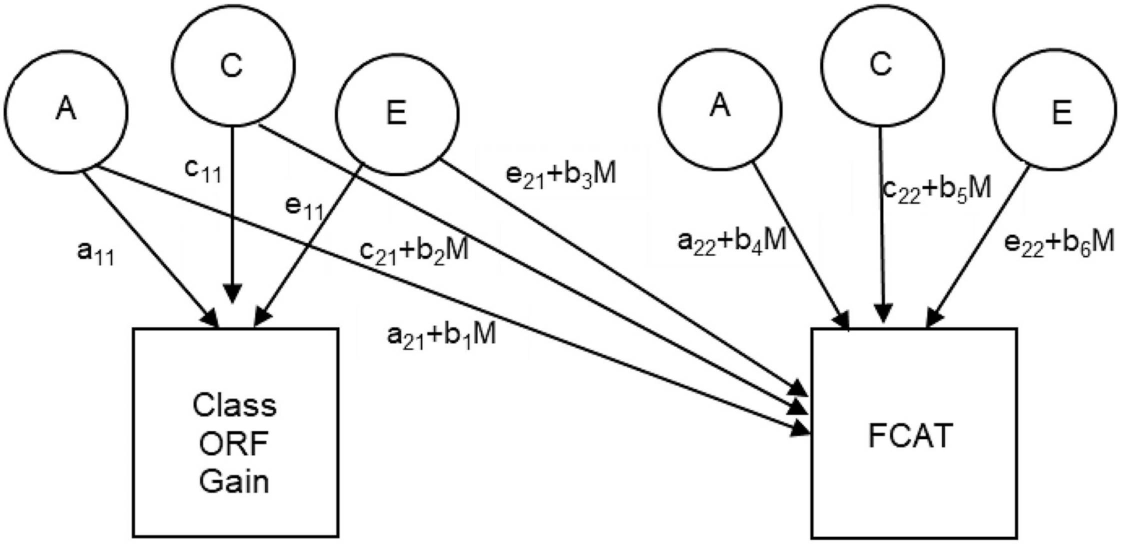

Figure 1.

Model testing moderation of additive genetic (A), shared environmental (C), and non-shared environmental (E) influences on FCAT (reading comprehension) by levels Class ORF Gain. The moderator (shown in the triangle) can influence variance that is common to it and the outcome as denoted with the subscript “C” (e.g., A_C) and/or the variance unique to the outcome as denoted by subscript “U” (e.g., A_U). *Note:* ORF = Oral Reading Fluency; FCAT = Florida Comprehensive Assessment Test in reading.

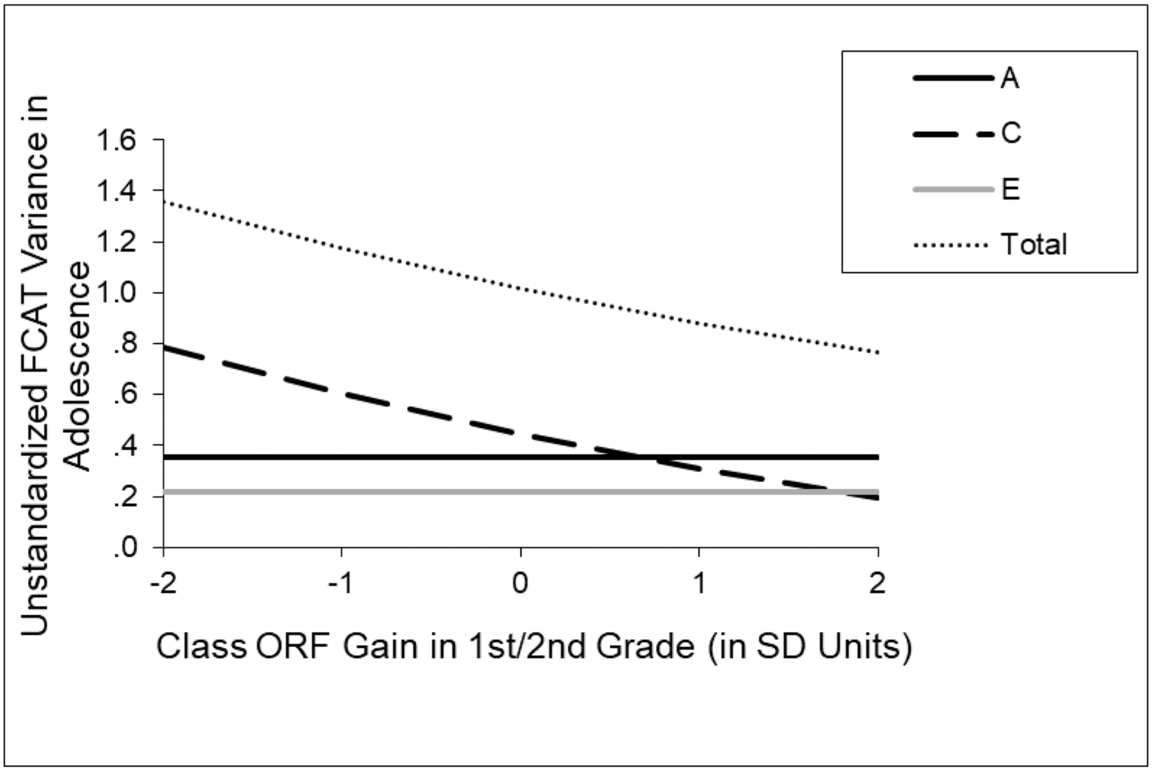

Figure 2.

Moderation of genetic and shared environmental effects on adolescent reading comprehension by level of quality in the early classroom environment. Total variance is shown along with unstandardized variance in additive genetic (A), shared environmental (C), and non-shared environmental (E) influences on FCAT (reading comprehension) by levels of Class ORF Gain. Based on the best-fitting model, the variance reflected in A and E includes only the variance unique to FCAT. The variance reflected in C includes the shared environmental influences common to Class ORF Gain and FCAT (39% of the total C shown) and shared environmental influences unique to FCAT (61% of the total C shown). For instance, at −2 *SD* units of Class ORF Gain, the variance in C is .79, of which .31 (.79 × .39) is common C and .48 (.79 × .61) is unique C. *Note:* ORF = Oral Reading Fluency; FCAT = Florida Comprehensive Assessment Test in reading.

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Mean Reading Comprehension (FCAT) by Level of Class ORF Gain and Zygosity

Class ORF Gain in 1st/2nd grade (in *SD* units)	FCAT in Adolescence									
	MZ					DZ				
	M (SD)	N	Min	Max	Skew	M (SD)	N	Min	Max	Skew
−2	211.67 (27.98)	6	175	246	−0.09	–	–	–	–	–
−1	236.74 (21.38)	70	185	294	0.17	227.78 (22.12)	103	175	282	0.16
0	239.97 (20.75)	325	167	296	−0.23	244.07 (20.90)	601	180	296	−0.12
1	253.67 (21.43)	48	211	302	0.05	256.45 (18.14)	101	218	302	0.44
2	259.95 (17.02)	21	231	296	0.28	258.68 (16.74)	40	211	302	−0.17
Overall	241.42 (21.87)	470	167	302	−0.14	244.26 (21.96)	845	175	302	−0.15

Note. N is number of individuals. ORF = Oral Reading Fluency; MZ = monozygotic; DZ = dizygotic; FCAT = Florida Comprehensive Assessment Test in reading.

Table 2.

Phenotypic and Twin Correlations by Level of Class ORF Gain

	FCAT in Adolescence				
Class ORF Gain in 1st/2nd grade (in *SD* units)	**Phenotypic**	**Intraclass**		**Cross-twin cross-trait**	
	(N)	**MZ (N)**	**DZ (N)**	**MZ (N)**	**DZ (N)**
−2	.82 (6)	*.67* (6)	-	*.43* (6)	-
−1	.23 (173)	.74 (70)	.66 (99)	*.18* (70)	.21 (102)
0	.23 (923)	.75 (316)	.60 (584)	.24 (324)	.23 (599)
1	*.11* (149)	.78 (48)	.18 (101)	*.25* (48)	*.04* (102)
2	*.07* (61)	.81 (21)	*.27* (39)	*.22* (21)	*.18* (40)
Overall	.37 (1312)	.77 (461)	.60 (823)	.36 (469)	.34 (843)

Note. Phenotypic correlation is between FCAT and Class ORF Gain scores. N is number of individuals. ORF = Oral Reading Fluency; MZ = monozygotic; DZ = dizygotic; FCAT = Florida Comprehensive Assessment Test in reading. All correlations were significant at $p < .05$ except those in italics.

Table 3.

Best-fitting Set of Models Testing Moderation of Influences on Adolescent Reading Comprehension by Early Classroom Reading Gains

Model Fitting	−2LL	df	AIC	Sample Size Adjusted BIC	$\Delta\chi^2$	Δdf	*p*	W_i
Full Moderation on All Common and All Unique Paths	6942.15	2849	1244.15	−1489.78				
No Moderation	6959.01	2855	1249.01	−1491.79	16.87	6	.01	<.01
Moderation on All Unique Paths	6950.49	2852	1246.49	−1490.83	8.34	3	.04	<.01
Moderation on All Common Paths	6948.92	2852	1244.92	−1491.61	6.78	3	.08	<.01
Moderation on Common C and Unique A, C, and E	6942.31	2851	1240.31	−1493.18	0.17	2	.92	.04
Moderation on Common C only	6949.64	2854	1241.64	−1494.74	7.49	5	.19	.02
Moderation on Common C and Unique C and E	6943.17	2852	1239.17	−1494.49	1.03	3	.79	.08
Moderation on Common C and Unique A and E	6949.03	2852	1245.03	−1491.56	6.89	3	.08	<.01
Moderation on Common C and Unique A and C	6942.32	2852	1238.32	−1494.91	0.18	3	.98	.12
Moderation on Common C and Unique E	6949.61	2853	1243.61	−1493.01	7.47	4	.11	.01
Moderation on Common C and Unique C	**6943.33**	**2853**	**1237.33**	**−1496.15**	**1.36**	**4**	**.85**	**.19**
Moderation on Common C and Unique A	6949.10	2853	1243.10	−1493.27	6.95	4	.14	.01

Note. $-2LL$ = −2 log-likelihood; AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; $\Delta\chi^2$ = chi-square change or the difference in $-2LL$ values between the full model and the given reduced model; Δdf = the difference in df for the chi-square change test; W_i = weighted AIC. Selected best-fitting model is in bold type. Analyses were conducted on standardized data, controlling for sex, age, and age-squared.