HHS Public Access

Author manuscript

Merrill Palmer Q (Wayne State Univ Press). Author manuscript; available in PMC 2017 July 01.

Published in final edited form as:

Merrill Palmer Q (Wayne State Univ Press). 2016 July ; 62(3): 233–251. doi:10.13110/merrpalmquar1982.62.3.0233.

Test anxiety and a high-stakes standardized reading comprehension test: A behavioral genetics perspective

Sarah G. Wood,
Florida State University

Sara A. Hart,
Florida State University

Callie W. Little, and
Florida State University

Beth M. Phillips
Florida State University

Abstract

Past research suggests that reading comprehension test performance does not rely solely on targeted cognitive processes such as word reading, but also on other non-target aspects such as test anxiety. Using a genetically sensitive design, we sought to understand the genetic and environmental etiology of the association between test anxiety and reading comprehension as measured by a high-stakes test. Mirroring the behavioral literature of test anxiety, three different dimensions of test anxiety were examined in relation to reading comprehension, namely intrusive thoughts, autonomic reactions, and off-task behaviors. Participants included 426 sets of twins from the Florida Twin Project on Reading. The results indicated test anxiety was negatively associated with reading comprehension test performance, specifically through common shared environmental influences. The significant contribution of test anxiety to reading comprehension on a high-stakes test supports the notion that non-targeted factors may be interfering with accurately assessing students' reading abilities.

Keywords

reading comprehension; twins; test anxiety; behavioral genetics

In the United States, there is an increasing demand for children to demonstrate their literacy proficiency, including reading comprehension, through high-stakes standardized tests, which are tests determining grade advancement or classroom placement (Afflerbach, 2005; Porter, McMaken, Hwang, & Yang, 2011). New educational policies such as the No Child Left Behind (NCLB) and Common Core legislation have called for the increased use of these tests as a means to measure educational learning, resulting in the introduction of high-stakes

Correspondence should be addressed to the second author at the Department of Psychology, 1107 W. Call Street, Florida State University, Tallahassee, FL 32306-4301. Phone: (850) 645-9693; FAX: (850) 644-7739. hart@psy.fsu.edu.
Sarah G. Wood, Department of Psychology; Sara A. Hart, Department of Psychology and Florida Center for Reading Research; Callie W. Little, Department of Psychology; Beth M. Phillips, Department of Psychology and Florida Center for Reading Research.

tests in more schools and at younger ages (Segool, Carlson, Goforth, von der Embse, & Barterian, 2013; Triplett & Barksdale, 2005; Wigfield & Eccles, 1989).

Reading comprehension is a complex task involving both bottom-up processes, such as decoding, and top-down processes, such as inference making (Cain & Oakhill, 2007; Snow, 2002). Reading comprehension becomes particularly important for school success during the “reading to learn” school years, commonly defined as grade 3 or 4 and beyond (Chall, 1996). At this time, classes begin to rely on students gaining knowledge in all subjects through the reading and understanding of text. To measure these students’ progress in school, particularly after third grade, standardized tests of reading comprehension have been developed and are, in many cases, state-mandated. Although this testing enables one to track students’ performance, it is not without complications.

Standardized tests of reading comprehension are designed to measure proficiency in reading comprehension, but other unintended situational conditions may impact test performance (Afflerbach, 2005; Haladyna, Haas, & Allison, 1998; von der Embse & Hasson, 2012). For example, cognitive components such as inattention (Huizink, van Beijsterveldt, Boomsma, & Bartels, 2011), or affective components such as low motivation (Paris, Lawton, Turner, & Roth, 1991), are commonly linked with poor test performance. In particular, there has been substantial research focused on test anxiety impacting testing outcomes (Haladyna & Downing, 2004; Segool et al., 2013). Test anxiety is a type of state anxiety (i.e., context-specific) specific to testing situations that impacts a student’s performance on the test, thus inhibiting the test score as an accurate reflection of academic knowledge and skill (Ergene, 2003). Between 10 – 40% of all students experience some level of test anxiety, with these rates varying, depending on factors such as gender, race and socio-economic status (McDonald, 2010; Segool et al., 2013; von der Embse et al., 2013). In general, the literature indicates a negative relation between test anxiety and academic test performance (Hembree, 1988; Seipp, 1991). Importantly, recent work has also highlighted the negative association of test anxiety with standardized tests of school achievement (Cassady & Johnson, 2002; Putwain, 2008; Segool et al., 2013; von der Embse & Hasson, 2012; von der Embse & Witmer, 2014).

In a study from the United Kingdom, results indicated a small but significant negative relation between test anxiety and achievement on the General Certificate of Secondary Education (GCSE) standardized test in secondary students, with test anxiety accounting for 7% of the variance in GCSE scores (Putwain, 2008). In the United States, elementary students were found to experience more test anxiety for state standardized tests than for classroom tests (Segool et al., 2013), with higher test anxiety levels associated with lower performance on high-stakes standardized tests, accounting for about 2 – 15% of the variance (von der Embse & Hasson, 2012; von der Embse & Witmer, 2014).

Cognitive load theory suggests that the executive control of working memory is limited, in that only a finite amount of resources are available to process new information (Sweller, 1988). It is thought that test anxiety can overwhelm the working memory system, making the processing of information necessary for reading comprehension inefficient and therefore more difficult (Cassady, 2004; Richards, French, Keogh, & Carter, 2000). Research in school-aged students indicates that children with higher test anxiety had lower total reading comprehension performance, although the exact nature and level of impact remains unclear (Gifford & Marston, 1966; Neville, Pfost, & Dobbs, 1967). For example, Gifford and Marston (1966) studied the effects of high and low test anxiety levels, reading rate, and task experience (pretest vs posttest) in a study of 4th grade boys. Although the study could not directly compare the test anxiety group's reading comprehension outcomes, the authors state that generally better reading comprehension is associated with lower test anxiety levels. Stronger support for this trend in children is found in Neville et al. (1967). Reading comprehension and vocabulary skills growth were measured in three groups of children differing in levels of test anxiety (high, medium and low). Children in the high test anxiety group had lower performance on standardized tests of reading comprehension than the children in the low anxiety group. Students in the medium level anxiety group performed better than both the low anxiety and high anxiety groups (Neville et al., 1967). Taken together, test anxiety in children is generally considered to be negatively related to reading comprehension performance.

To move beyond simply measuring the phenotypic association between test anxiety and reading comprehension, it is possible to examine the extent to which there are genetic and environmental influences between the two, using a twin design (Neale & Cardon, 1992). A twin design enables the investigation of the relative contribution of genetic influences, or heritability (i.e., additive genetic influences inherited from your parents; h^2), shared environmental influences (i.e., environmental effects that serve to make siblings more similar; c^2), and nonshared environmental influences (i.e., environmental effects that serve to make siblings less similar; e^2) on any given construct independently, and importantly here, the genetic and environmental influences in common among different constructs. Specifically, by comparing monozygotic twins, who share 100% of their genetic influences, to dizygotic twins, who share on average 50% of their genetic influences, the genetic and environmental contributions of a skill or behavior of interest such as test anxiety and reading comprehension can be estimated. Additionally, the role of genetic and/or environmental influences underlying the association between test anxiety and reading comprehension performance can be quantified, allowing for a better understanding of risk factors and narrowing of possible sources of confounds. For example, if genetic influences were found to explain the association between test anxiety and reading comprehension, then genetic risk factors in common between the two, perhaps related to cognitive skill, would be indicated. If shared environmental influences were found to explain the same association, then environmental risk factors, perhaps related to the shared school-environment, would be indicated. Importantly, it is typically not possible to determine the exact source of the genetic and/or shared environmental influences, but the information gathered from the twin method allows for a narrowing of the possible sources. If nonshared environmental influences were found to explain the association between test anxiety and reading comprehension, than this might suggest a causal relation between the two (Turkheimer & Harden, 2014). This is because many possible sources of confounds, typically shared between members of a twin family, are controlled for by the twin model. Therefore, in total, the twin model allows for a richer description of the association between test anxiety and reading comprehension than previously available by phenotypic correlations only.

The previous behavioral genetics literature has not directly examined test anxiety, although work has focused on state anxiety. State anxiety is defined as anxiety induced by specific situations only (Spielberger, 1996). Test anxiety is a subtype of state anxiety. This previous work indicates that state anxiety has moderate shared environmental (c^2 =.30 – .36) and large non-shared environmental (e^2= .59 – .84) influences, with negligible genetic factors (Lau, Eley, & Stevenson, 2006; Legrand, McGue, & Iacono, 1999). Behavioral genetics work focusing on reading comprehension has indicated both significant genetic and environmental factors. Estimates of heritability have been small to large, ranging from .32 to .82 (Betjemann, Keenan, Olson, & DeFries, 2011; Betjemann et al., 2008; Hart, Petrill, & Kamp Dush, 2010; Keenan, Betjemann, Wadsworth, DeFries, & Olson, 2006; Logan et al., 2013; Petrill, Deater-Deckard, Schatschneider, & Davis, 2007). Additionally, there are low to moderate shared environmental influences on reading comprehension, ranging from .01 to .47 (Betjemann et al., 2008; Logan et al., 2013; Petrill et al., 2007), and low to moderate non-shared environmental influences, ranging from .13 to .31 (Betjemann et al., 2011; Byrne et al., 2009; Keenan et al., 2006).

The cognitive load theory and the previous literature suggest that high test anxiety is associated with lower performance on tests of reading comprehension. Given the nature of high-stakes testing, which has become the standard metric to measure successful reading achievement for school children (i.e., failure of the test can result in grade retention), it is likely that test anxiety plays an even more important role in these tests. Using a twin sample in Florida, we sought to examine the etiology of this association by examining the genetic and environmental contributions to the association between test anxiety and reading comprehension, as measured by a high-stakes test. Based on previous work, we hypothesized that test anxiety would indicate both shared and nonshared environmental influences, and that reading comprehension would indicate genetic and environmental influences. As there has been no research examining the genetic and environmental association between test anxiety and reading comprehension, we did not have a hypothesis for this relation.

Methods

Participants

Participants were drawn from the Florida Twin Project on Reading, Behavior and Environment, a representative ongoing cross-sequential twin project in the state of Florida (for more information see (Taylor, Hart, Mikolajewski, & Schatschneider, 2013). The project obtained monitoring and achievement data for reading from the statewide educational database, Progress Monitoring and Reporting Network (PMRN), as well as additional data concerning twin behavior and environment via a parent and twin (for children at least 9 years old) questionnaire mailed directly to families beginning in 2010. For the present study, twins who were in grades 3 through 7 during the 2010–2011 school year were analyzed, representing all data available (mean age = 11.82yrs, SD = 1.08yrs). This resulted in a final sample of 213 monozygotic (MZ; 118 female-female pairs) and 213 same-sex dizygotic (DZ; 106 female-female pairs) twin pairs with complete data. According to parent report, twins were 54% Caucasian, 22% Hispanic, 16% African American, 5% mixed race, and 3%

other or unknown. Zygosity of the twin pairs was determined via a parental five-item questionnaire on physical similarity (Lykken et al., 1990).

Procedure and measures

Twins' test anxiety was measured by a self-questionnaire collected on all twins 9 years old and older during the summer of 2010. Reading comprehension data was collected by trained administrators as part of statewide achievement testing required by normal school attendance, and test scores were uploaded into the PMRN via a web-based data collection system. PMRN data collection windows are determined by the Florida Department of Education and local school districts; for the present study all data were collected in spring 2011, the time period most immediately following the questionnaire data collection. These procedures were conducted in accordance with the American Psychological Association (APA) ethical principles for human subject research, including informed consent and assent.

Test anxiety—Children's Test Anxiety Scale (CTAS; Wren & Benson, 2004) was used to measure test anxiety. The 30-item scale has three dimensions, "thoughts", "off-task behaviors", and "autonomic reactions" measured on a 4-point Likert scale (1=almost never, 4=almost always). "Thoughts" is defined as test irrelevant thoughts, self-critical thoughts and other types of worry. "Off-task behaviors" is defined by distracting behaviors or nervous habits invoked by test anxiety. "Autonomic reactions" is defined as somatic changes related to test anxiety (e.g., sweating, upset stomach). Students were asked to respond to each question with the directions "how you think, feel, or act, when taking a test". Mean scores for each dimension were used, as well as a mean score for the whole measure (i.e., test anxiety total score), and in all cases, higher scores indicate higher test anxiety. Reliability for each was sufficient, thoughts alpha = .91, off-task behavior alpha = .82, autonomic reactions alpha = .84, and test anxiety total score alpha = .94.

FCAT—Reading comprehension performance was measured by the Florida Comprehensive Achievement Test (FCAT) 2.0 Reading test. This criterion-referenced test was group-administered to all Florida students every spring. During this test, students answered short or long format multiple-choice items depending on the content of the passage. Reliability for FCAT Reading Comprehension from item response theory (IRT) ranges from 0.90 in 3rd grade to 0.92 in Grades 5–12 (Foorman & Petscher, 2010). FCAT 2.0 standard scores, ranging from 100–500, were used in the analyses.

Analyses

To investigate the relations between test anxiety and reading comprehension, phenotypic correlations and genetically-sensitive Cholesky decomposition models were used (DeFries & Fulker, 1985; Neale, Boker, Xie, & Maes, 2006). Phenotypic correlations, analyzed using SAS 9.4, were used to establish initial relationships between test anxiety and reading comprehension. Then, four bivariate Cholesky decomposition models were estimated, three separate models for each of the test anxiety dimensions and reading comprehension and one model for the full test anxiety measure and reading comprehension. More specifically, we examined the bivariate relation between thoughts and FCAT, that between off-task behaviors and FCAT, that between autonomic reactions and FCAT, and that between test anxiety total score and FCAT. Each bivariate Cholesky decomposition partitions the variance on a variable, and covariance shared between two variables, into two sets of biometric factors (see Figure 1). The first set of biometric factors represent the additive genetic, shared environmental variance (environmental influences that make siblings more similar) and nonshared environmental variance (environmental influences that make siblings less similar, plus error) shared between the test anxiety dimension and reading comprehension (A_1, C_1, and E_1 respectively). The second set of biometric factors represents the additive genetic, shared environmental and nonshared environmental variance unique to reading comprehension alone (A_2, C_2, and E_2 respectively), after accounting for the covariance between test anxiety and reading comprehension. After descriptive statistics and phenotypic correlations were calculated, all analyses were done with z-scored data corrected for age, age x age, gender and age x gender in Mx (McGue & Bouchard, 1984; Neale & Maes, 2013). Prior to the final modeling, the test anxiety dimension mean scores were reverse scored, so the variance between test anxiety and reading comprehension was positive, to assist in modeling.

Results

Table 1 presents descriptive statistics for the dimensions of test anxiety and FCAT scores. All skewness values were within the normal range, although the autonomic reaction dimension was positively skewed in the moderately non-normal range (<2). Pearson correlations between all measures are listed in Table 2. Correlations between the dimensions of test anxiety were large and significant (r = .55 – .73 p<.001), and as could be expected the dimensions were significantly correlated with the test anxiety total score (r = .86 – .89 p<. 001). Correlations between the measures of test anxiety and FCAT were low to moderate and significant (r = −.29 – −.09, p<.05).

Intraclass Correlations

Next, intraclass correlations by zygosity were analyzed to provide an initial descriptive estimate of heritability, shared environmental and nonshared environmental influences on each dimension of test anxiety and FCAT scores (see Table 3). Genetic influences were inferred for all measures because monozygotic (MZ) twin correlations were higher (r_{MZ} = . 37–.71) than those for dizygotic (DZ) twins (r_{DZ} = .32–.61) in all cases. Additionally, shared environmental influences are indicated because MZ correlations were less than twice DZ correlations. Finally, non-shared environmental influences (along with error) were indicated as MZ correlations were less than one (Neale & Cardon, 1992).

Bivariate Cholesky Analyses

Three bivariate Cholesky decomposition models were used to examine the genetic and environmental contributions to the variance and covariance between each dimension of test anxiety and FCAT scores. Table 4 presents the univariate variance components for each variable and the path estimates from the four bivariate Cholesky decomposition models. Briefly, univariate results across the models indicated that the test anxiety dimensions of thoughts indicated significant genetic (h^2), shared environmental (c^2) and nonshared environmental (e^2) influences. Off-task behavior indicated significant genetic and nonshared environmental influences only and autonomic reactions and test anxiety total score indicated significant shared and nonshared environmental influences only. FCAT scores indicated significant genetic, shared environmental and nonshared environmental influences.

Turning towards the bivariate path estimates, when test anxiety was defined as the dimension of thoughts, results revealed significant genetic influences on thoughts and FCAT alone (path estimates of .59 and .74, respectively), but no significant genetic overlap between the two. On the other hand, there were significant shared environmental influences in common between thoughts and FCAT, with no additional shared environmental influences left on FCAT after accounting for the common effect. Looking at the relation between the off-task behavior dimension of test anxiety and FCAT, there were no significant common influences between the two. However, both indicated significant unique genetic and nonshared environmental influences. Next, in the last dimension model, where test anxiety was represented by the autonomic reactions dimension the only significant genetic influences were unique to FCAT (path estimate of .74). Also, there were significant common shared environmental influences between the two measures, accounting for all of the shared environmental influences on FCAT, with no additional shared environmental influences on FCAT after accounting for the common effect. For this model, there were also significant common nonshared environmental influences between autonomic reactions and FCAT.

Lastly, when test anxiety was defined as the total score results showed the only significant overlapping influence between test anxiety and FCAT was shared environment (path estimate of .38). Significant independent influences on test anxiety were found for genetic, shared and non-shared environmental influences (path estimates of .56, .43 and .70 respectively). In this last model independent influences on FCAT reflect the previous models findings of significant independent genetic and non-shared environmental influences (path estimates of .74 and .50 respectively).

Discussion

Through a twin approach, both genetic and environmental influences on and between test anxiety and reading comprehension were explored. Three different dimensions of test anxiety, namely thoughts, off-task behaviors, and autonomic reactions, were investigated in relation to a high-stakes reading comprehension achievement test.

As a first step, phenotypic correlations were examined, as the findings concerning the relation between test anxiety and reading comprehension did not previously include standardized measures of reading comprehension. The present results indicated that all three dimensions of test anxiety were negatively associated with FCAT scores, supporting the hypothesis that test anxiety in general is significantly related to scores on the high-stakes FCAT reading test. Interestingly, the correlation of off-task behaviors with reading comprehension was statistically smaller than the other two test anxiety dimensions with reading comprehension (Meng et al., 1992)[1] . Although on the surface it seems counter-intuitive that off-task behaviors are not as strongly associated with reading comprehension test scores as the other measures of test anxiety (i.e., being off-task would seemingly be a problem for testing), previous work with this scale has suggested that this dimension is not as closely related to the overall test anxiety construct. Specifically, factor analysis of the scale has indicated this dimension has the lowest factor loading of the three dimensions (Putwain & Daniels, 2010). The present result supports this previous finding, indicating that this measure of off-task behaviors is simply not a hallmark measure of test anxiety compared to the other two dimensions.

Turning towards the genetically-sensitive modeling, contrary to our original hypothesis based on the state anxiety literature, which indicated environmental influences only, when examining the univariate models, both genetic and environmental factors were indicated for test anxiety. It may be the case that test anxiety taps into different aspects of personality which are more heritable, such as perfectionism (Tozzi et al., 2004), resulting in greater genetic influences than other anxiety traits (Legrand et al., 1999; Lau et al., 2006). Interestingly, the results indicated that despite the literature suggesting that there are categorically different dimensions of test anxiety (e.g., Wren & Benson, 2004), the underlying etiology of these supposedly different dimensions is remarkably similar. The univariate results for the FCAT are in line with previous reports from this twin project (Hart et al., 2013), as well as the previous literature on reading comprehension (Keenan et al., 2006; Logan et al., 2013).

The main research question of this study was to examine the genetic and environmental influences underlying the association of test anxiety and reading achievement as measured by a high-stakes achievement test. The literature indicates that test anxiety could potentially be a reason for why some children do poorly on these sorts of tests, despite proficient reading skills. Given the implications for doing poorly on these mandated tests, for the student, the teacher, and the school, it is important to understand the nature of the shared relation. The results across the bivariate models were fairly consistent. If there was a shared etiology between test anxiety and FCAT, it was through the shared environment, indicating that there were environmental confounds to the association between the two. This result was seen for the dimensions of thoughts and autonomic reactions, but not for off-task behavior, likely reflecting the low overall relation between off-task behavior and FCAT ($r = -.09$). Importantly, when the test anxiety measure was used as a whole (rather than the different dimensions), the results supported this same finding of shared environmental overlap. It has been hypothesized that anxiety symptoms can overwhelm the working memory system, making it difficult to process the information necessary to be successful on a test (Coy, OBrien, Tabaczynski, Northern, & Carels, 2011). At first glance, it seems counter-intuitive that the working memory system may be activated through common shared environmental influences. However, given that test anxiety is a state-specific anxiety, it often occurs during the testing situation, and related events, only. The FCAT is an environmental event that is shared across twin pairs because they take the test at the same time with similarly aged peers in the same school. FCAT testing is typically a stressful time in the school year, with school administration, teachers and students all impacted by the results of the testing. This can

[1]The correlation contrast test was used to determine if the correlation of each of the test anxiety measures with FCAT were statistically different from each other. Results indicated the only significant differences were that the correlations between thoughts and FCAT ($r = -.28$), and autonomic reactions and FCAT ($r = -.29$) were significantly greater than off-task behaviors and FCAT ($r = -.09$; z = −5.53, $p = .00$ and z = −5.54, $p = .00$, respectively).

result in the broader school climate being more anxious, which has been linked with increasing test anxiety in children within the school (Segool, von der Embse, Mata, & Gallant, 2014). This more anxious school climate may be contributing to the shared environmental influences measured here. Interestingly, we suspect but can't test more specifically that children in third grade are more likely than children in other grades to feel anxiety due to the high stakes nature of the third grade reading test where passing is required to advance to the next grade.

Limitations and future work

There are several limitations of the study worth noting. First, the test anxiety literature suggests there may be sex differences in the expression of test anxiety (McDonald, 2010). Unfortunately, we have only same-sex dizygotic twins that do not allow for the full sex-limitation model to be explored, and most importantly, our sample size was too small once separated by sex to detect sex differences. However, if we were able to examine for sex effects, we may have found a differential relation between test anxiety and reading by sex. Along these same lines, it has been suggested that there is a developmental pattern indicating that the relation between test anxiety and test performance may increase across the elementary school years, stabilizing around grade 5 (Hembree 1988). For example, Hill and Sarason (1996) found that the correlation between test anxiety and test performance was lower for boys in 3^{rd} grade (boys $r = -.18$, girls $r = -.28$) than in 5^{th} grade (boys $r = -.34$, girls $r = -.24$). It may be the case that there are developmental patterns that we were unable to analyze because the available sample size per grade is not high enough to do individual twin analyses[2] .

Another limitation to note is that the test anxiety measure was not given at the same time of FCAT testing, so although there is a shared etiology between test anxiety and FCAT, it is likely that the children were thinking of all testing situations and not the FCAT specifically. Finally, another important limitation is that these models do not account for the direction of the effect. It is unknown if test anxiety leads to poor reading test scores, or if children who struggle with reading comprehension develop test anxiety because of the poor prior performance. Alternatively, it may be the case that the association between test anxiety and reading comprehension test performance may be due to another variable not included here, such as generalized anxiety (in this sample, generalized anxiety disorder and test anxiety are correlated at $r = .23$). Knowing the nature of the causative relation is important for determining the implications of these findings on the growing momentum towards using high-stakes testing in schools.

As high-stakes testing becomes more generally adopted by the public education system, it is more important than ever to understand how factors such as stress and anxiety may interfere with a student's performance. These results support test anxiety's negative impact on reading comprehension, specifically through common shared environmental influences. The

[2]A phenotypic test of this possibility was examined. Four multiple regressions were run for each measure of test anxiety, with test anxiety and with the interaction of test anxiety and grade entered as predictors of FCAT scores. In all cases, the interaction term was a nonsignificant predictor of reading comprehension (thoughts and FCAT, t-value = −.33, $p = .74$; off-task behaviors and FCAT, t-value = −1.08, $p = .28$; autonomic reactions and FCAT, t-value = −.52, $p = .60$; and test anxiety total score and FCAT, t-value = −0.92, $p = .36$).

significant contribution of test anxiety to reading comprehension on a high-stakes test supports the notion that test anxiety may be interfering with accurately assessing students' reading abilities. This has important implications for the use of high-stakes testing in schools, and the subsequent decisions from these tests such as automatic grade retention.

Acknowledgments

This research was supported, in part, by a grant from the National Institute of Child Health and Human Development (P50 HD052120). Views expressed herein are those of the authors and have neither been reviewed nor approved by the granting agencies. We wish to thank the twins and their families for their participation in this research.

References

Afflerbach P. National Reading Conference Policy Brief: High Stakes Testing and Reading Assessment. Journal of Literacy Research. 2005; 37(2):151–162. DOI: 10.1207/s15548430jlr3702_2

Betjemann RS, Keenan JM, Olson RK, DeFries JC. Choice of Reading Comprehension Test Influences the Outcomes of Genetic Analyses. Scientific Studies of Reading. 2011; 15(4):363–382. DOI: 10.1080/10888438.2010.493965 [PubMed: 21804757]

Betjemann RS, Willcutt EG, Olson RK, Keenan JM, DeFries JC, Wadsworth SJ. Word reading and reading comprehension: stability, overlap and independence. Reading and Writing. 2008; 21(5):539–558. DOI: 10.1007/s11145-007-9076-8

Byrne B, Coventry WL, Olson RK, Samuelsson S, Corley R, Willcutt EG, et al. Genetic and environmental influences on aspects of literacy and language in early childhood: Continuity and change from preschool to Grade 2. Journal of Neurolinguistics. 2009; 22(3):219–236. DOI: 10.1016/j.jneuroling.2008.09.003 [PubMed: 20161176]

Cain, K., Oakhill, J. Children's Comprehension Problems in Oral and Written Language: A Cognitive Perspective. Cain, K., Oakhill, J., editors. New York: The Guilford Press; 2007.

Cassady JC. The impact of cognitive test anxiety on text comprehension and recall in the absence of external evaluative pressure. Applied Cognitive Psychology. 2004; 18(3):311–325. DOI: 10.1002/acp.968

Cassady JC, Johnson RE. Cognitive Test Anxiety and Academic Performance. Contemporary Educational Psychology. 2002; 27(2):270–295. DOI: 10.1006/ceps.2001.1094

Chall, JS. Stages of reading development. 2. Orlando: Harcourt Brace; 1996.

Coy B, OBrien WH, Tabaczynski T, Northern J, Carels R. Associations between evaluation anxiety, cognitive interference and performance on working memory tasks. Applied Cognitive Psychology. 2011; 25(5):823–832. DOI: 10.1002/acp.1765

DeFries JC, Fulker DW. Multiple regression analysis of twin data. Behavior Genetics. 1985; 15(5):467–473. DOI: 10.1007/BF01066239 [PubMed: 4074272]

Ergene T. Effective Interventions on Test Anxiety Reduction A Meta-Analysis. School Psychology International. 2003; 24(3):313–328. DOI: 10.1177/01430343030243004

Gifford EM, Marston AR. Test Anxiety, Reading Rate, and Task Experience. The Journal of Educational Research. 1966; doi: 10.2307/27531725

Haladyna TM, Downing SM. Construct-Irrelevant Variance in High-Stakes Testing. Educational Measurement: Issues and Practice. 2004; 23(1):17–27. DOI: 10.1111/j.1745-3992.2004.tb00149.x

Haladyna T, Haas N, Allison J. Continuing Tensions in Standardized Testing. Childhood Education. 1998; 74(5):262–273.

Hart SA, Petrill SA, Kamp Dush CM. Genetic Influences on Language, Reading, and Mathematics Skills in a National Sample: An Analysis Using the National Longitudinal Survey of Youth. Language, Speech, and Hearing Services in Schools. 2010; 41(1):118–128. DOI: 10.1044/0161-1461(2009/08-0052)

Hembree R. Correlates, Causes, Effects, and Treatment of Test Anxiety. Review of Educational Research. 1988; 58(1):47–77. DOI: 10.3102/00346543058001047

Hill K, Sarason S. The relation of test anxiety and defensiveness to test and school performance over the elementary-school years: A further longitudinal study. Monographs of the Society for Research in Child Development. 1996; 31(2):1–75. DOI: 10.2307/1165770

Huizink AC, van Beijsterveldt CEM, Boomsma DI, Bartels M. A genetic study on attention problems and academic skills: results of a longitudinal study in twins. Journal of the Canadian Academy of Child and Adolescent Psychiatry = Journal De l'Académie Canadienne De Psychiatrie De L'Enfant Et De L'adolescent. 2011; 20(1):22–34. DOI: 10.1177/0956797610386617

Keenan JM, Betjemann RS, Wadsworth SJ, DeFries JC, Olson RK. Genetic and environmental influences on reading and listening comprehension. Journal of Research in Reading. 2006; 29(1): 75–91.

Lau JYF, Eley TC, Stevenson J. Examining the State-Trait Anxiety Relationship: A Behavioural Genetic Approach. Journal of Abnormal Child Psychology. 2006; 34(1):18–26. DOI: 10.1007/s10802-005-9006-7

Legrand LN, McGue M, Iacono WG. A Twin Study of State and Trait Anxiety in Childhood and Adolescence. Journal of Child Psychology and Psychiatry. 1999; 40(6):953–958. DOI: 10.1111/1469-7610.00512 [PubMed: 10509889]

Logan JAR, Hart SA, Cutting L, Deater-Deckard K, Schatschneider C, Petrill S. Reading Development in Young Children: Genetic and Environmental Influences. Child Development. 2013; 84(6):2131–2144. DOI: 10.1111/cdev.12104 [PubMed: 23574275]

McDonald AS. The Prevalence and Effects of Test Anxiety in School Children. Educational Psychology. 2010; doi: 10.1080/01443410020019867

McGue M, Bouchard TJ Jr. Adjustment of twin data for the effects of age and sex. Behavior Genetics. 1984; 14(4):325–343. DOI: 10.1007/BF01080045 [PubMed: 6542356]

Meng XL, Rosenthal R, Rubin DB. Comparing correlated correlation coefficients. Psychological bulletin. 1992; 111(1):172.

Neale, MC., Maes, HHM. Methodology for Genetic Studies of Twins and Families. Dordrecht: Kluwer Academic Publishers B.V; 2013. p. 1-308.

Neale MC, Boker SM, Xie G, Maes HH. Mx Statistical Modeling. Richmond. 2006

Neale, M., Cardon, LR. Methodology for genetic studies of twins and families. 1992.

Neville D, Pfost P, Dobbs V. The Relationship between Test Anxiety and Silent Reading Gain. American Educational Research Journal. 1967; 4(1):45–50. DOI: 10.3102/00028312004001045

Paris SG, Lawton TA, Turner JC, Roth JL. A Developmental Perspective on Standardized Achievement Testing. Educational Researcher. 1991; 20(5):12–20. DOI: 10.3102/0013189X020005012

Petrill SA, Deater-Deckard K, Schatschneider C, Davis C. Environmental influences on reading-related outcomes: an adoption study. Infant and Child Development. 2007; 16(2):171–191. DOI: 10.1002/icd.476

Porter A, McMaken J, Hwang J, Yang R. Common Core Standards The New U.S. Intended Curriculum. Educational Researcher. 2011; 40(3):103–116. DOI: 10.3102/0013189X11405038

Putwain DW. Test anxiety and GCSE performance: the effect of gender and socio-economic background. Educational Psychology in Practice Theory, Research and Practice in Educational Psychology. 2008; 24(4):319–334. DOI: 10.1080/02667360802488765

Putwain DW, Daniels RA. Is the relationship between competence beliefs and test anxiety influenced by goal orientation? Learning and Individual Differences. 2010; 20(1):8–13. DOI: 10.1016/j.lindif.2009.10.006

Richards A, French CC, Keogh E, Carter C. Test-Anxiety, inferential reasoning and working memory load. Anxiety, Stress & Coping. 2000; 13(1):87–109. DOI: 10.1080/10615800008248335

Segool NK, Carlson JS, Goforth AN, von der Embse N, Barterian JA. Heightened test anxiety among young children: elementary school students' anxious responses to high-stakes testing. Psychology in the Schools. 2013; 50(5):489–499. DOI: 10.1002/pits.21689

Segool NK, von der Embse N, Mata AD, Gallant J. Cognitive Behavioral Model of Test Anxiety in a High-Stakes Context: An Exploratory Study. School Mental Health. 2014; 6(1):50–61. DOI: 10.1007/s12310-013-9111-7

Seipp B. Anxiety and academic performance: A meta-analysis of findings. Anxiety Research. 1991; 4(1):27–41. DOI: 10.1080/08917779108248762

Snow, CE. Reading for understanding: Toward a research and development program in reading comprehension. Arlington, VA: RAND; 2002. p. 1-174.

Spielberger, CD. Anxiety and behavior. Oxford, England: Academic Press; 1996.

Taylor JE, Hart SA, Mikolajewski AJ, Schatschneider C. An Update on the Florida State Twin Registry. Twin Research and Human Genetics. 2013; 16(01):471–475. DOI: 10.1017/thg.2012.74 [PubMed: 23067863]

Tozzi F, Aggen SH, Neale BM, Anderson CB, Mazzeo SE, Neale MC, Bulik CM. The Structure of Perfectionism: A Twin Study. Behavior Genetics. 2004; 34(5):483–494. DOI: 10.1023/B:BEGE.0000038486.47219.76 [PubMed: 15319571]

Triplett CF, Barksdale MA. Third through Sixth Graders' Perceptions of High-Stakes Testing. Journal of Literacy Research. 2005; 37(2):237–260. DOI: 10.1207/s15548430jlr3702_5

Turkheimer, E., Harden, KP. Behavior Genetic Research Methods. In: Reis, HT., Judd, CM., editors. Handbook of Research Methods in Social and Personality Psychology. Cambridge University Press; 2014. p. 159-187.

von der Embse N, Hasson R. Test Anxiety and High-Stakes Test Performance Between School Settings: Implications for Educators. Preventing School Failure: Alternative Education for Children and Youth. 2012; 56(3):180–187. DOI: 10.1080/1045988X.2011.633285

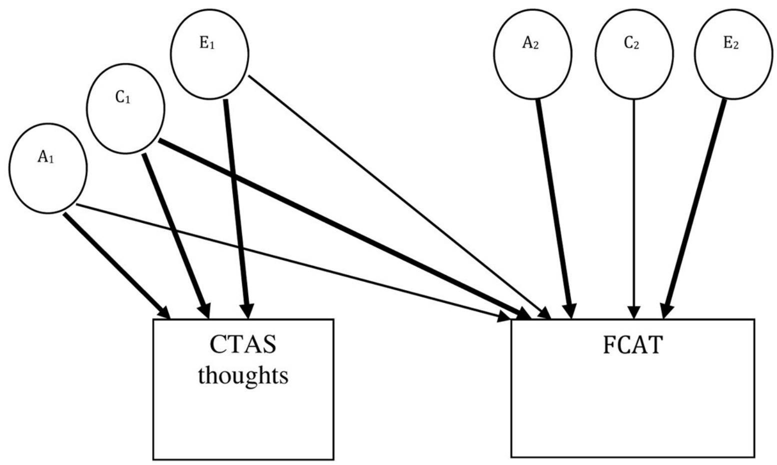
von der Embse N, Witmer SE. High-Stakes Accountability: Student Anxiety and Large-Scale Testing. Journal of Applied School Psychology. 2014; 30(2):132–156. DOI: 10.1080/15377903.2014.888529

von der Embse N, Barterian J, Segool N. Test Anxiety Interventions for Children and Adolescents: A Systematic Review of Treatment Studies from 2000–2010. Psychology in the Schools. 2013; 50(1):57–71. DOI: 10.1002/pits.21660

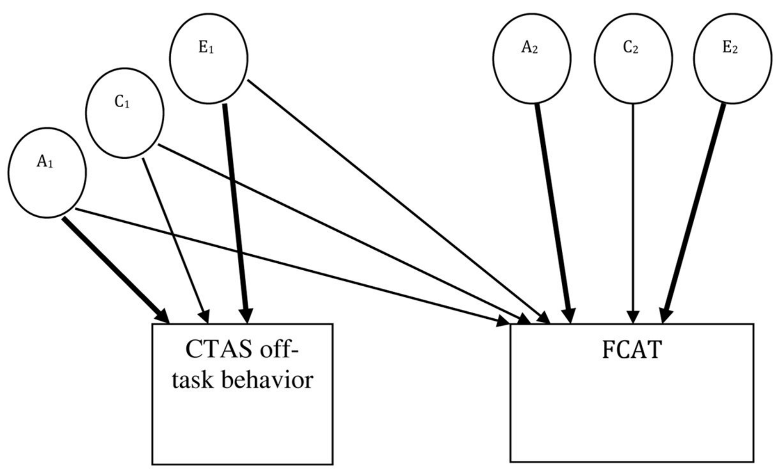
Wigfield A, Eccles JS. Test Anxiety in Elementary and Secondary School Students. 1989; 24(2):159–183. Dx.Doi. org. DOI: 10.1207/s15326985ep2402_3

Wren DG, Benson J. Measuring test anxiety in children: Scale development and internal construct validation. Anxiety, Stress & Coping. 2004; 17(3):227–240. DOI: 10.1080/10615800412331292606

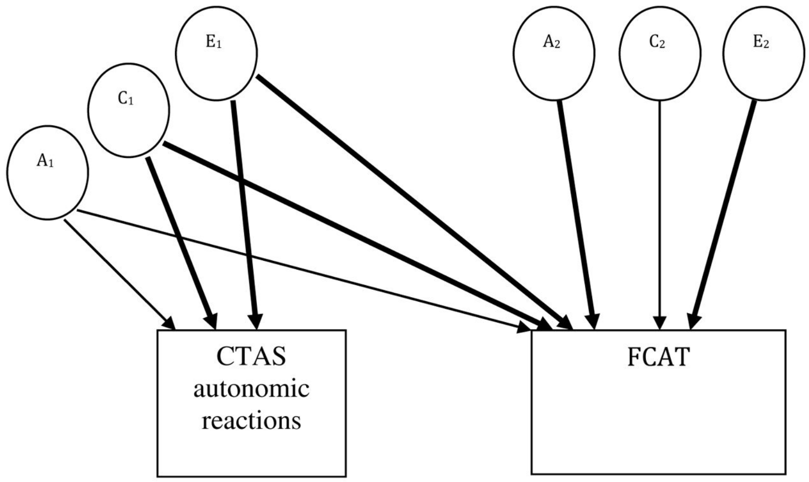
A


E_1
C_1
A_1
A_2
C_2
E_2
CTAS thoughts
FCAT

B


E_1
C_1
A_1
A_2
C_2
E_2
CTAS off-task behavior
FCAT

C



D

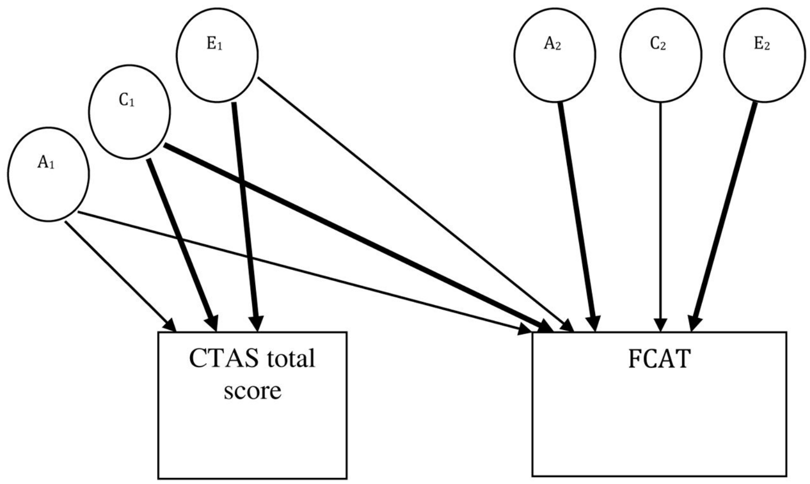


Figure 1.

Bivariate Cholesky modeling of test anxiety and the Florida Comprehensive Achievement Test (FCAT). Each of the three bivariate Cholesky models (panels A–C) contained a test anxiety dimension (thoughts, off-task behavior, and autonomic reactions, respectively) and FCAT. Additionally, a fourth bivariate Cholesky model for the full test anxiety measure and reading comprehension was included (panel D). In the model, variance and covariance were decomposed into additive genetic (A), shared environmental (C), and nonshared environmental (E) factors. Significant pathways are noted by bold arrows. CTAS, Children's Test Anxiety Scale.

Table 1

Means, standard deviations (SD), minimums, maximums and skewness for grades 3 through 7 Test Anxiety and FCAT scores.

Variable	Mean	SD	Minimum	Maximum	Skew	*n*
CTAS Thoughts[a]	2.11	.66	1.00	4.00	.61	853
CTAS Off-Task Behavior[a]	1.92	.65	1.00	4.00	.83	855
CTAS Autonomic Reactions[a]	1.58	.54	1.00	4.00	1.44	849
CTAS Total Score[a]	1.90	.54	1.00	4.00	.87	849
FCAT	327.69	59.81	100.00	500.00	−.21	543

[a]Children's Test Anxiety Scale (CTAS)

Table 2

Phenotypic correlations for combined grades 3 through 7 Test Anxiety and FCAT scores (n in parentheses).

Variable	CTAS Thoughts[a]	CTAS Off-Task Behavior[a]	CTAS Autonomic[a]	CTAS Total Score[a]	FCAT
CTAS Thoughts[a]	1.00 (853)				
CTAS Off-Task Behavior[a]	.59** (852)	1.00 (855)			
CTAS Autonomic Reactions[a]	.73** (847)	.55* (847)	1.00 (849)		
CTAS Total Score[a]	.89** (853)	.83** (855)	.86** (849)	1.00 (857)	
FCAT	−.28** (541)	−.09* (543)	−.29* (536)	−.24** (544)	1.00 (744)

Note.

* $p < .05$

** $p < .001$

[a] Children's Test Anxiety Scale (CTAS)

Table 3

Intraclass correlations for grades 3 through 7 Test Anxiety and FCAT scores for monozygotic (MZ) and dizygotic (DZ) twin pairs (n in parentheses).

	Twin intra-class correlations	
Variable	**MZ**	**DZ**
CTAS Thoughts[a]	.47* (419)	.34* (415)
CTAS Off-Task Behavior[a]	.49* (420)	.38* (419)
CTAS Autonomic Reactions[a]	.37* (418)	.32* (409)
CTAS Total Score[a]	.46* (422)	.37* (419)
FCAT	.71* (253)	.61* (259)

Note.

* $p < .0001$

[a] Children's Test Anxiety Scale (CTAS)

Table 4

Univariate variance components of additive genetic (h^2), shared environmental (c^2) and nonshared environmental (e^2) estimates of test anxiety and FCAT, as well as bivariate Choleksy modeling path estimate results of genetic (A), shared environmental (C) and nonshared environmental influences between test anxiety and FCAT [with 95% confidence intervals].

Variable	Univariate estimates	Shared influences between test anxiety and FCAT	Independent influences on FCAT
	h^2	A_1	A_2
CTAS Thoughts[a]	.35* [.09–.58]	.59* [.29–.76]	
FCAT	.55* [.43–.67]	.06 [.00–.38]	.74* [.63–.81]
	c^2	C_1	C_2
CTAS Thoughts[a]	.15* [.001–.37]	.39* [.04–.61]	
FCAT	.20* [.09–.31]	.45* [.03–.55]	.00 [.00–.48]
	e^2	E_1	E_2
CTAS Thoughts[a]	.49* [.41–.60]	.70* [.64–.77]	
FCAT	.25* [.22–.29]	.04 [.00–.11]	.50* [.47–.54]
	h^2	A_1	A_2
CTAS Off-Task Behavior[a]	.35* [.08–.63]	.59* [.28–.79]	
FCAT	.55* [.43–.68]	.00 [.00–.21]	.74* [.66–.82]
	c^2	C_1	C_2
CTAS Off-Task Behavior[a]	.18 [.00–.41]	.43 [.00–.64]	
FCAT	.19* [.09–.30]	.14 [.00–.52]	.41 [.00–.53]
	e^2	E_1	E_2
CTAS Off-Task Behavior[a]	.46* [.38–.56]	.68* [.62–.75]	
FCAT	.25* [.22–.29]	.00 [.00–.05]	.50* [.47–.54]
	h^2	A_1	A_2
CTAS Autonomic Reactions[a]	.27 [.00–.48]	.52 [.00–.69]	
FCAT	.55* [.44–.67]	.00 [.00–.30]	.74* [.66–.82]
	c^2	C_1	C_2
CTAS Autonomic Reactions[a]	.15* [.01–.37]	.39* [.09–.61]	
FCAT	.19* [.09–.30]	.44* [.10–.54]	.00 [.00–.46]

Variable	Univariate estimates	Shared influences between test anxiety and FCAT	Independent influences on FCAT
	e^2	E_1	E_2
CTAS Autonomic Reactions[a]	.56* [.47–.68]	.75* [.69–.83]	
FCAT	.25* [.22–.29]	.08* [.01–.15]	.49* [.46–.53]
	h^2	A_1	A_2
CTAS Total Score[a]	.32* [.03–.55]	.56 [.18–.74]	
FCAT	.55* [.43–.67]	.00 [.00–.31]	.74* [.66–.82]
	c^2	C_1	C_2
CTAS Total Score[a]	.18* [.01–.41]	.43* [.11–.64]	
FCAT	.19* [.09–.30]	.38* [.03–.54]	.21 [.00–.49]
	e^2	E_1	E_2
CTAS Total Score[a]	.49* [.41–.60]	.70* [.64–.77]	
FCAT	.25* [.22–.29]	.04 [.00–.10]	.50* [.47–.54]

Note.

* indicates significance based on confidence intervals not bounding zero.

[a] Children's Test Anxiety Scale (CTAS)

Author Manuscript

Author Manuscript

Author Manuscript