Psychology in the Schools, Vol. 52(4), 2015
View this article online at wileyonlinelibrary.com/journal/pits

DOI: 10.1002/pits.21827

GENDER DIFFERENCES IN ACHIEVEMENT IN A LARGE, NATIONALLY REPRESENTATIVE SAMPLE OF CHILDREN AND ADOLESCENTS

CAROLINE SCHEIBER, MATTHEW R. REYNOLDS, DANIEL B. HAJOVSKY, AND ALAN S. KAUFMAN

Alliant International University, University of Kansas, and Yale University

The purpose of this study was to investigate developmental gender differences in academic achievement areas, with the primary focus on writing, using the child and adolescent portion (ages 6–21 years) of the *Kaufman Test of Educational Achievement—Second Edition, Brief Form,* norming sample ($N = 1{,}574$). Path analytic models with gender, parent education, age, age^2, and gender-by-age moderation as predictors of reading, writing, and math were used to test for gender differences and for the influence of development on these differences. A small but consistent advantage was identified for females in reading. No gender differences were detected in math. The most important results of the present study pertain to a gender gap in writing in favor of females that increased as a function of age. Male students are at greater risk for writing failure than are females. © 2015 Wiley Periodicals, Inc.

Despite generating a great deal of media attention, gender differences in cognitive abilities and academic achievement are often much smaller than one is led to believe (Hyde, 2005). For example, it is often publicized that males score higher in math than females, but in reality such findings are sometimes either small or nonexistent or depend on general ability level, age, or on the specific math skill assessed (e.g., Lachance & Mazzocco, 2006; McGraw, Lubienski, & Strutchens, 2006). Alternatively, one gender difference in academic achievement that is rarely publicized and almost never emphasized, despite a warning that was published about 20 years ago in *Science* (Hedges & Nowell, 1995), is a female advantage in writing.

If school psychologists are asked to consult with school-based intervention teams with the goal of delivering large-scale academic interventions efficiently, it is important to recognize whether and when demographic characteristics such as gender may put a group of students at risk of falling behind in school (e.g., males in writing). The primary purpose of this investigation was, therefore, to examine developmental gender differences in reading, math, and writing in a large and nationally stratified sample of children and adolescents on an individually administered test of achievement. Of specific interest was the study of gender differences in writing, as measured by the *Kaufman Test of Educational Achievement—Second Edition, Brief Form* (KTEA-II Brief; Kaufman & Kaufman, 2005).

Gender Differences in Academic Achievement

Writing Achievement. Researchers have reported gender gaps in writing in favor of females across different populations and designs. For example, females have demonstrated higher writing scores within special populations, such as in individuals with dyslexia (Berninger, Nielsen, Abbott, Wijsman, & Raskind, 2008); on group achievement tests such as the National Assessment of Educational Progress (NAEP; Persky, Danee, & Jin, 2003; Salahu-Din, Persky, & Miller, 2008; U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2011a); and on measures collected longitudinally (Martin & Hoover, 1987).

Correspondence to: Caroline Scheiber, 14055 Pomerado Road, San Diego, CA 92064 USA. E-mail: carolinescheiber@gmail.com

The corresponding author address has been changed on 23rd june 2015.

One valuable piece of evidence of gender differences comes from research using large representative samples of children and adolescents who are administered achievement tests by individual examiners. Camarata and Woodcock (2006) conducted a comprehensive study of gender differences in cognitive abilities and academic achievement using data from stratified norming samples of the Woodcock–Johnson (WJ) tests. In that study, females demonstrated a substantial writing advantage (.33–.43 *SD*) over males across the 5–79 years age range on three versions of the WJ. Nonetheless, studies on gender differences in writing with nationally representative samples using individually administered writing tests are rare. Kaufman, Kaufman, Liu, and Johnson (2009) conducted such a study with adults, but only Camarata and Woodcock reported comparable data for Grades 1–12. The measurement of writing can vary substantially across different tests, so a study with an individually administered test other than the WJ test is needed for school-aged children.

One additional finding related to gender differences in writing is a developmental trend related to age. For example, one study showed that a female advantage in writing increased with age during the school years (Malecki & Jewell, 2003). Nevertheless, those findings were based on scores from short and timed curriculum-based measures of writing. Studies of gender differences using timed tests may confound differences in writing ability with differences in writing speed (Camarata & Woodcock, 2006). New studies are needed to test for possible gender differences in writing that are moderated by age with the potential speed confound removed from the scores.

To summarize, although female advantages in writing have been documented, multiple sources of evidence across multiple designs, samples, and instruments provide better evidence of empirical realities with regard to constructs (Lykken, 1968). Facts are only established after a thorough process of constructive replication. If gender differences in writing generalize across the well-designed WJ tests to the well-designed yet substantially different Kaufman tests, then such gender differences are more likely related to *constructs* rather than to a specific instrument. Such findings are important for school psychologists and educators, as they allow for intervention strategies targeted for males who may be at relative risk for writing failure.

Reading and Math Achievement. Females have demonstrated a small and sometimes moderate advantage in reading among school-aged children, which has also been shown across different countries (e.g., Mullis, Martin, Kennedy, & Foy, 2007). Nationally, in 2011 females outperformed males in fourth ($d = .20$) and eighth grades ($d = .26$) on the NAEP reading assessments (U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2011b). Female advantages in reading, however, are not unequivocal. Kaufman et al. (2009) did not find significant gender differences in reading in their nationally representative sample of *adults*, suggesting that if any early differences exist, they may disappear by adulthood. Moreover, Camarata and Woodcock (2006) did not find any gender differences in reading in their sample across the age range of 5 to 79 on untimed WJ reading tests.

Findings related to gender differences in math have also been inconsistent. Results of a meta-analysis showed that males begin to outperform females in high school on mathematics problem solving, and that the discrepancy increases with age (Hyde, Fennema, & Lamon, 1990). Similarly, Camarata and Woodcock (2006) found that males in high school and adult cohorts (except for college students) performed better than females on WJ math achievement tests; a finding that was similar to those found with a nationally representative sample of adults using the Kaufman tests (Kaufman et al., 2009). Camarata and Woodcock found no gender differences in math in the youngest samples in their research—kindergarten, elementary school, or middle school. Longitudinal research has also shown that gender differences in math skills are either too inconsistent to be considered important or are nonexistent in elementary school, supporting the cross-sectional findings from Camarata and Woodcock (Lachance & Mazzocco, 2006). In addition to age, gender differences in math also seem to depend on the type of skill assessed. For example, although males tend to outperform females on tests that require math concepts, females tend to outperform males on math computation (Gallagher et al., 2000; Martin & Hoover, 1987). Research with NAEP data also seems to support findings related to school-age children and skill assessed. For example, McGraw et al. (2006) analyzed NAEP data between 1990 and 2003 and found that the male advantages in math were generally small or nonexistent, depending on the math skill assessed, but were pronounced for those students that scored at the higher end of the distribution.

Overall, gender differences in reading and math have been found to be small and sometimes moderate, but are often too inconsistent to draw definitive conclusions during the *school-age years*. The focus of the current study was on gender differences in writing. But, in this study it was important to utilize measures of reading and math, along with measures of writing, in the same sample of children because it permitted comparisons of gender differences across achievement areas while holding the sample constant. Specifically, findings related to gender differences in writing would be strengthened, for example, if they existed in the presence of small or inconsistent differences in general reading and math during the school-age years.

Study Purpose

The current study was the first to use the child and adolescent portion of the KTEA-II Brief Form (Kaufman & Kaufman, 2005) to investigate gender differences. The use of the norming data from this instrument had several advantages and findings could be used to contribute to the current literature in several ways: (a) The test is individually administered so that optimal performance may be obtained for each person in the sample (e.g., Thorndike, 2005); (b) the sample was stratified on important background variables, including ethnicity, geographic region, and parents' educational attainment according to 2001 U.S. Census data; (c) the KTEA-II Brief has multiple achievement domains so gender differences may be compared across domains within the same sample; (d) the data can be used to extend previous findings of the adult portion (ages 22–90) of the standardization samples from the KTEA-II Brief (Kaufman et al., 2009)[1] to children and adolescents (ages 6–21)—providing a representative sample across the lifespan—and to explore whether the differences depend on the age of the individuals assessed and whether findings with curriculum-based writing measurements of an increasing female advantage in writing generalize to other writing measures, specifically those without a timed component (e.g., Malecki & Jewell, 2003); and (e) perhaps most importantly, the KTEA-II Brief measures writing in a novel and interactive way that differs substantially from previous measures. The writing task is administered in the form of writing and editing a newsletter. A female advantage on a measure of writing that differs from the WJ tests and from curriculum-based writing measures in its measurement of writing provides strong evidence that it is related to the construct of writing.

We answered two questions in this research, with a focus on writing:

1. Are there gender differences on the KTEA-II Brief in reading, math, and writing in children and adolescents (ages 6–21), after statistically controlling for parent education and linear and quadratic age effects?
2. Does age moderate the influence of gender on reading, math, and writing after statistically controlling for parent education and linear and quadratic age effects?

[1] Kaufman et al. (2009) investigated the relationship of educational attainment and gender to academic skills and intelligence (fluid and crystallized) in adulthood.

Method

Participants

The total number of examinees (ages 4.5–90 years) in the KTEA-II Brief norm sample constituted 2,495 individuals. The sampling procedure consisted of two stages. First, demographic information was collected on a large pool of potential examinees. Second, random sampling procedures were applied in choosing the examinees that ultimately became part of the study. The participants were stratified on key background variables (i.e., sex, ethnicity, geographic region, and educational level) and the sample was nationally representative of children and adolescents in kindergarten through 12th grade according to 2001 U.S. Census (Kaufman & Kaufman, 2005). Participants selected for the present study were in Grades 1 or higher to ensure at least one year of academic instruction; adults of ages 22–90 years were excluded. There were 1,574 participants: 793 females and 781 males. The sample ranged in age from 6 to 21 years ($M = 11.87$, $SD = 4.27$), and comprised 257 Black (16.3%), 274 Latino/a (17.4%), 960 White (61%), 58 Asian (3.7%), and 14 Native American (0.9%) participants. Eleven participants identified with "other" ethnicity (0.7%). Parent education was the educational attainment of the examinee's mother or female guardian; the father's or male guardian's education was used only when information on the mother was unavailable (i.e., 1% of the cases). The breakdown of parent education and coding scheme is as follows: 1 = Grade 11 or less ($n = 213$; 13.5%), 2 = high school graduate or GED ($n = 501$; 31.8%), 3 = some college/tech. certification/associate degree ($n = 492$; 31.3%), and 4 = bachelor's degree or higher ($n = 368$; 23.4%).[2] The mean parent education of the sample was 2.64 ($SD = .98$).

Measures

KTEA-II Brief. The KTEA-II Brief (Kaufman & Kaufman, 2005) is a succinct individually administered test of educational achievement. The KTEA-II Brief subtests are reported as age-standardized scores with a maximum score range from 40 to 160 ($M = 100$, $SD = 15$). Content validity was established for each area of achievement for both the KTEA-II Brief and the Comprehensive Forms, "by using a variety of sources, including literature reviews, surveys given to special-education teachers and school psychologists, and recommendations from experts in curriculum and special-education assessment" (Kaufman & Kaufman, 2005, p. 34). Furthermore, subtests correlate moderately (.50–.70) with each other, which suggests independent yet interrelated constructs (Urbina, 2014). Correlations of KTEA-II Brief subtests with reading, math, and writing subtests on other well-known achievement tests such as the WJ III provide strong support of its convergent and discriminant validity (Kaufman & Kaufman, 2005, Tables 6.9–6.19). Finally, the KTEA-II Brief shows age-related decline and growth patterns consistent with the literature (Kaufman, Johnson, & Liu, 2008).

The KTEA-II Brief measures three achievement areas: reading, writing, and math. The reading test consists of recognition (automatic word recognition and phonetic decoding) and comprehension. The assessment of both skills is consistent with the domain of reading. Recognition items are set up to "ensure that the subtest measures word recognition (reading vocabulary) more than decoding ability" (Kaufman & Kaufman, 2005, p. 3). Comprehension items primarily assess the ability to read a passage and provide answers to literal or inferential questions. To compute internal-consistency reliabilities of the test, split-half methods were used. The subtests were divided into parallel halves by using an odd–even split and then, with the aim to achieve balance of item difficulty and content, the composition of each half was adjusted accordingly (Kaufman & Kaufman, 2005, p. 55). Split-half reliability estimates for reading averaged .94 for children and adolescents in Grades 1–12 and

[2] For more-detailed demographic information of the sample, please contact the first author.

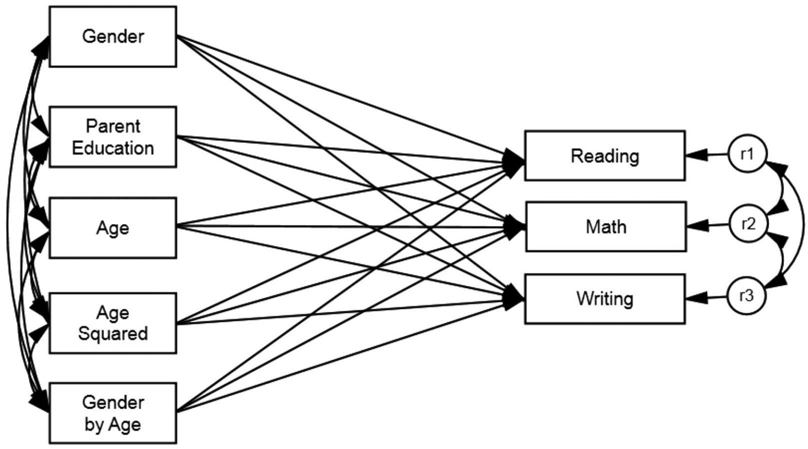

FIGURE 1. Initial path model for tests of gender differences in reading, math, and writing, controlling for parent education, gender × age moderation, and linear and quadratic age effects.

was .96 for ages 19–20; the test–retest coefficient (1-month interval) was .95 for Grades 1–2, .91 for Grades 4–7, and .91 for ages 15–21 years. The reading composite correlated strongly (.85) with the reading test of the WJ III (Kaufman & Kaufman, 2005).

The math test consists of computation skill and application categories. In the computation part, the individual is asked to solve mathematical problems and write them down. In the application part, the individual is asked to respond orally to real-life mathematical application questions. Split-half reliability estimates for math averaged .90 for children and adolescents in Grades 1–12 and was .93 for ages 19–20 years; the test–retest coefficient (1-month interval) was .84 for Grades 1–2, .89 for Grades 4–7, and .91 for ages 15–21 years. The math test correlates strongly (.70) with the WJ III math cluster (Kaufman & Kaufman, 2005).

The writing test is composed of writing and editing a newsletter. This writing format was chosen because it allows examinees to express their ideas in writing, provides context for writing, and offers an interactive writing experience. The individual is required to communicate ideas in writing without involving reading skills or reasoning ability or creativity, which makes this test a purer measure of writing. Split-half reliability estimates for writing averaged .86 for children and adolescents in Grades 1–12 and was .91 for ages 19–20 years; all items from any particular passage were always included on the same "half-test" to ensure experimental independence of each half-test. The test–retest coefficient (1-month interval) was .87 for Grades 1–2, .82 for Grades 4–7, and .81 for ages 15–21 years. The writing composite correlates strongly (.71) with the writing portion of the WJ III (Kaufman & Kaufman, 2005).

Analytic Steps

We utilized a just-identified path model with reading, math, and writing scores as outcome variables to answer the research questions (see Figure 1). This model allowed for the simultaneous estimation of all the predictor effects on multiple outcome variables rather than performing a series of separate moderated multiple regressions for each outcome variable. The model allowed for the inclusion of all data whether complete or incomplete, and all of the variables, and thus a more straightforward parsimonious interpretation of the findings.[3] We performed all analyses using Amos software version 20 (Arbuckle, 1995–2011).

Gender, parent education, age, age^2, and gender $\times$ age were included as predictor variables. Because nobody in the sample had a standard score of zero, prior to analysis, we centered the continuous predictor variables (i.e., age and parent education) to provide more meaningful interpretations of the findings (Aiken & West, 1991). We centered the age variable at age 12, which was close to the mean age of the sample. The mean of the age-centered variable therefore was 0, which represented age 12. We multiplied that centered variable by gender (females = 0 and males = 1) to create a gender-by-age cross-product variable. When the cross-product variable was included in the model along with gender and age, it was interpreted as a moderator effect, with its unstandardized coefficient representing the gender difference in age-achievement slopes. Next, we squared the age-centered variable to create an additional variable (age^2). This variable was also included in the model to protect against spurious moderator effects related to nonlinear age trends (Lubinski & Humphreys, 1990). Last, we centered the parent education variable at 2 (2 = graduated from high school) so that parent education of 0 indicated that the parent graduated from high school.

The path model is shown in Figure 1. The three outcome variables, reading, math, and writing, are located on the right side of the figure. We covaried the residual variances associated with the outcome variables. These covariances represented the shared variance among the three outcome variables that was not explained by the predictors (i.e., it was unrealistic to assume the variables in the model accounted for all of the shared variance among achievement tests). Located on the left side of Figure 1 are the gender, parent education, age, age^2, and gender $\times$ age predictor variables, which were all intercorrelated.

Nonstatistically significant gender $\times$ age and age^2 effects from the initial model were removed in subsequent models so that gender differences in its constituent effects could be interpreted as main effects. Gender and age effects in models with statistically significant moderators were interpreted as conditional effects. We performed post hoc probing in the presence of moderated effects (Cohen, Cohen, West, & Aiken, 2003; Keith, 2006).

Results

Univariate analysis of gender differences in the reading, writing, and math scores for the whole sample indicated statistically significant mean differences in reading and writing (see Table 1). Females, on average, scored higher in both areas, with a small effect size for reading and moderate-to-large effect size for writing. This analysis, however, did not include moderator or additional control variables.

Path Models

Seven estimates from the just-identified path model with gender, parent education, age, age^2, and gender-by-age as predictors of the three achievement outcomes were statistically significant ($p < .01$). We deleted four nonstatistically significant (i.e., based on the standard errors) moderator and quadratic effects. Model fit did not degrade when these were deleted ($\Delta\chi^2$ [4] = 4.21, p = .90), further substantiating that they were not statistically significant. This model, with statistically significant effects in bold, is shown in Figure 2. The unstandardized path estimates are shown in the Figure 2 and are subsequently interpreted with regard to each achievement variable. The numbers

[3]Only one case was missing one reading score. All cases were included and maximum-likelihood estimation procedures were used.

Table 1
Means and Standard Deviations by Gender for the Whole Sample, and Cohen's d Effect Sizes for Mean Differences

	Males		Females		Effect size
KTEA-II brief composite ($N = 1{,}574$)	*M*	*SD*	*M*	*SD*	Cohen's *d*
Reading[a]	98.46	14.70	101.11	14.45	−.18*
Math	101.01	16.01	100.90	15.09	−.01
Writing	97.63	14.53	103.55	14.82	−.40*

Note. KTEA-II Brief composite have a mean of 100 and an SD of 15. Minus sign with Cohen's d indicates females scored higher.
[a]One participant was missing a reading composite score.
$^{*}p < .05$.

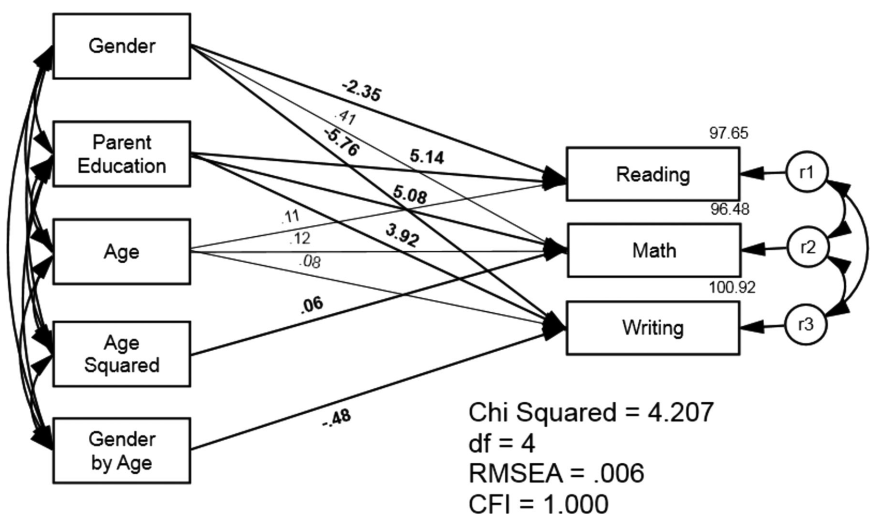

FIGURE 2. Gender, age, age^2, parent education, and gender-by-age effects on reading, math, and writing. All estimates are in unstandardized units. Nonstatistically significant age^2 and gender-by-age effects were deleted from this model. Path estimates from age and gender that are not bolded were not statistically significant. Intercepts are located on the upper-right corner of the achievement variables, and represent the predicted achievement score for females who were 12 years old and whose parents graduated from high school.

located on the upper right of the achievement variables in Figure 2 represent the intercepts, or the predicted achievement score for 12-year-old females whose parent graduated high school.

Reading

The combination of predictor variables explained 13% of the total variance in reading scores. There were no statistically significant effects of age^2 or gender-by-age on reading. We deleted those paths from the model (see Figure 2). There was not a statistically significant main effect of age on reading.[4] Statistically significant estimates for parent education and gender and associated confidence intervals are shown in Table 2. Parent education had a statistically significant effect on reading, after controlling for gender and age. A one-unit increase (e.g., from parent with less

[4]This finding is expected because age-standardized scores were used.

Table 2
Salient Parent Education, Gender, Age, Age^2, and Gender-by-Age Moderation Effects on KTEA-II Brief Reading, Math, and Writing Scores

			95% CI	
	b	*SE*	Lower	Upper
Gender				
Reading[a]	−2.35	0.68	−3.68	−1.02
Writing[b]	−5.76	0.71	−7.15	−4.37
Parent education				
Reading	5.14	0.35	4.45	5.83
Math	5.08	0.38	4.34	5.82
Writing	3.92	0.36	3.21	4.63
Age^2				
Math[c]	0.06	0.02	0.03	0.09
Gender × age				
Writing	−0.48	0.12	−0.72	−0.24
Age for males				
Writing[d]	−0.40	0.10	−0.60	−0.20

Note. All effects were statistically significant at $p < .01$.
[a]Main effect.
[b]Conditional mean difference because of the statistically significant moderation. That is, mean gender difference in writing scores for students who are 12 years old and whose parent education was 12 years.
[c]Linear coefficient associated with age was not statistically significant.
[d]Simple slope for males, not shown in the Figure 2.

than high school to high school education and so forth) in parent education was associated with a 5.14-point increase in predicted reading score ($\beta = .35$).[5]

There was a statistically significant main effect of gender on reading, controlling for parent education. The unstandardized estimate was −2.35, indicating that on average males scored 2.35 standard score points lower than females on reading. As shown on the upper-right side of the reading variable in Figure 2, the intercept was 97.65, which was the predicted reading score for a 12-year-old female whose parents graduated from high school. The predicted reading score for her male counterpart matched on the other variables was 95.30. The gender difference in reading was consistent across parent education and age levels.

Math

The combination of variables explained 11% of the total variance in math scores. There was no statistically significant gender-by-age effect found for math. We deleted this path from the model (Figure 2). There was not a statistically significant gender effect on math.

Parent education had a statistically significant effect on math scores. Statistically controlling for the other variables in the model, a one-unit increase in parent education was associated with 5.08-point increase in predicted math scores ($\beta = .32$). The quadratic effect, age^2, was statistically significant. The estimate was positive, indicating a concave shape of math scores with age controlling

[5]The parent education variable was treated as a continuous variable, although it was coarsely categorized. We also ran models with three dummy-coded parent education variables rather than as a single variable. Because the findings related to parent education were consistent across levels of education and it was not the focus of the study, we presented the findings with the single parent education variable for ease of exposition.

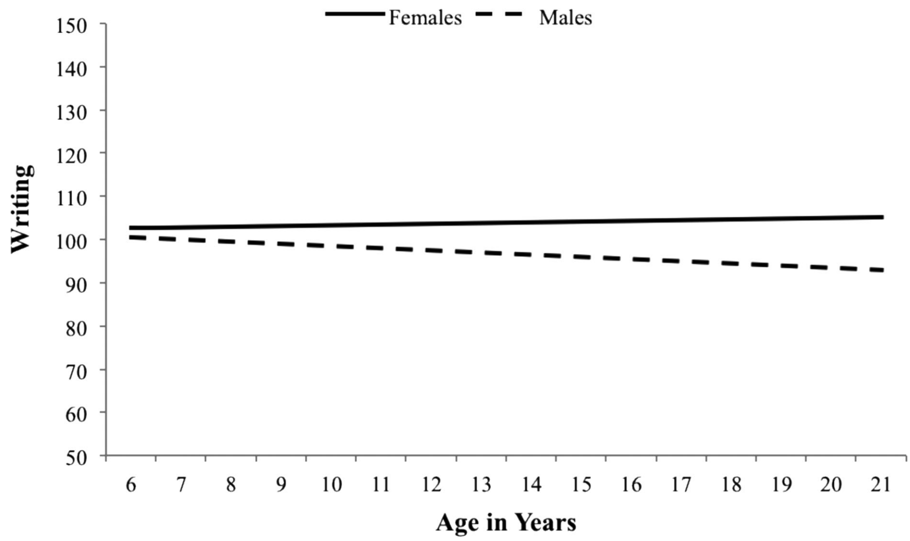

FIGURE 3. Age and writing relations plotted by gender.

for other variables in the model. The estimate associated with age to math was not statistically significant, so there was not a positive or negative trend across age. Although it was important to control for nonlinear trends, because age-standardized scores were used, the finding has little substantive value.

Writing

The combination of variables explained 12% of the total variance in writing scores. There was no statistically significant effect related to either age or age^2. Parent education had a statistically significant effect on writing scores. Statistically controlling for the other variables in the model, a one-unit increase in parent education was associated with 3.92-point increase in predicted writing scores ($\beta = .26$).

There was a statistically significant gender-by-age moderation effect for writing. All estimates related to gender and age, including those shown in Figure 2, were interpreted in the context of the moderated effect. To better understand the nature of the moderation, we plotted writing scores by age for males and females (Figure 3). A small writing advantage for females was evident at age 6, but that advantage increased with age.

The intercept, shown in Figure 2, represents the predicted writing score (100.92) for a 12-year-old female whose parents graduated high school. The path from gender to writing ($b = -5.76$) is the female–male difference in predicted writing scores for 12-year-olds whose parents graduated high school (i.e., males predicted score = 95.16). This effect, however, is a conditional effect. That is, as indicated by the statistically significant moderation effect, the writing difference depended on age.

Post hoc probing showed that the effect of age on writing for females (i.e., simple slope) was not statistically significantly different from zero ($b = 0.08$, $z[1,574] = .757$, $p = .449$). Females who differ from each other by a year in age differ by .08 in their writing scores, which was not statistically significant.

The difference in slopes from age to writing for males and females is represented by the path from gender-by-age to writing ($b = -0.48$). This value was statistically significantly different from zero, indicating moderation (Table 2). In addition to testing whether the slopes were different from each other, we tested whether the simple slope for males (that was associated with age and writing) was significantly different from zero. To obtain the correct standard errors for this test, we estimated another model that was equivalent to the previous model in terms of fit. For this model, however, we changed the gender variable so that males were coded as 0 and females were coded as 1 and we created a new gender-by-age cross-product. We replaced the gender variable and the gender-by-age cross-product variable from the previous model with these two new variables. The effect of age on writing for males was statistically significant ($b = -0.40$, $z[1{,}574] = -3.93$, $p < .001$). The predicted writing standard score for males decreased by a little less than a half of a point each year (indicating an increasing male and female difference in writing with age). This negative slope is shown in Figure 3.

Because the slopes between males and females, on average differed by .48 points, the female advantage in writing increased, on average, by about a half of a point per year. At age 6, females had about a three-point advantage in predicted writing scores; by age 12, the predicted advantage increased to about six points; and by age 20, the predicted advantage was approximately ten points, which is about a .75 *SD* advantage.

Discussion

The purpose of the present study was to investigate gender differences in reading, math, and writing in a stratified sample of children and adolescents across ages 6–21 years using a well-designed, individually administered measure of achievement. Parent education and linear and quadratic age effects were controlled whereas gender-by-age moderation was tested. Our findings revealed a small advantage for females in reading that was consistent across the age range. Females also showed an advantage in writing, but the gap increased as a function of age, resulting in a substantial gender difference in favor of females in the oldest age groups. Males and females did not differ in math achievement. Last, parent education had statistically significant effects on all three of the achievement areas.

Gender Differences

Reading and Math. A small but stable female advantage in reading was detected. These results are consistent with some recent findings (Mullis et al., 2007) but contrast with Camarata and Woodcock (2006), who found no gender differences in their sample across the lifespan. Kaufman et al. (2009) analyzed the adult portion of the standardization sample of the KTEA-II Brief, and found no gender differences in reading at ages 22–90 years. Taken together, small gender differences in reading on the KTEA-II Brief exist for children and adolescents, but these small differences seem to disappear for adults.

No gender differences in math achievement were detected in the present sample, but Kaufman et al. (2009) did identify small gender differences in favor of adult males (.28 *SD*) at ages 22–90 years. Small gender differences in math on the KTEA-II Brief only emerged, however, in adult cohorts. This pattern is very similar to the pattern shown with the WJ tests, except Camarata and Woodcock (2006) found that a male advantage in math appeared in high school and beyond (except for within a college sample), as compared to at age 22.

Writing. Females demonstrated a considerable writing advantage over males in a representative sample of typically developing children and adolescents. The magnitude of the difference increased with age. After calculating Cohen's *d* effect sizes in different age groups, females exhibited a small advantage starting as early as ages 6–10 years ($d = .25$). The small difference then transformed into a moderate-to-large effect by ages 15–18 years ($d = .64$). The gender differences appear to stabilize in adolescence and remain through adulthood; for example, Kaufman et al. (2009) found a consistent moderate-to-large writing advantage in their adult sample in favor of females (.59 *SD*). Taken together with findings from research with other nationally representative samples and well-designed individually administered writing tests (e.g., Camarata & Woodcock, 2006), the current findings provide support that gender differences in writing are due to the construct and are not instrument specific.

Implications

Given the findings in this study and others, researchers, school psychologists, and educators may need to pay special attention to gender differences in writing. The current research demonstrates that males are more vulnerable to writing failure than are females.

The exploration of the causes of gender differences in writing is still in its early stages of development; however, some researchers have studied gender differences in writing and have offered potential explanations. For example, some studies found differences in the lateralization of language between males and females (Shaywitz et al., 1995), and others found differences in orthographical or phonological skills (Berninger et al., 2008; Berninger & Fuller, 1992; Schneider & Näslund, 1999). Females also show a clear advantage in processing speed abilities (Camarata & Woodcock, 2006), and processing speed abilities have been shown to predict writing skills (Floyd, McGrew, & Evans, 2008). Such differences may play an important role in explaining writing differences.

Alternatively, research has identified environmental variables that could potentially explain the widening of the gender writing gap. For example, it has been shown that teachers often interact less with students who struggle with writing and tend to focus their attention on other students (Graham & Harris, 1997). If males are more likely to struggle with writing early on, then it may be that they are also less likely to be instructed by teachers (and in addition the subtle rejection can easily result in a loss of hope and confidence in writing ability). To counteract this possible effect, educators could focus additional attention to those students who struggle (and who are more likely to be male), using interventions that have been found to improve writing proficiency, such as providing explicit writing instruction in the classroom (De La Paz & Graham, 2002). This additional attention may also allow students to experience their successes in the process of developing effective writing skills.

School psychologists may play an important role in taking early educational action to improve the development of writing for males. One potential way to improve writing for males is to create instruction delivery processes that are especially appealing to male students, but the research in this area is currently extremely limited. King and Gurian (2006) summarized specific "boy-friendly teaching strategies" (p. 56) that have shown to be effective in improving males' writing abilities in one elementary school. For example, teachers have found it helpful to split the class into two groups and have each group arrange cards to form grammatically correct sentences. Similarly, other writing interventions that were used in a collaborative group work setting and those that stimulated visual–spatial skills have been found to be of particular benefit to males (Hidi, Berndorff, & Ainley, 2002). Furthermore, writing is believed to be more of a female-oriented task and this gender orientation is positively correlated with females' writing proficiency (e.g., Pajares, Valiante, & Cheong, 2007). Therefore, it may be helpful to dispel this belief. For example, exposure to male role models who speak in front of the class about their own writing and literacy may help reduce gender-oriented beliefs. Moreover, having males choose topics for writing may help stimulate their interests (Goiran-Bevelhimer, 2008).

Whereas research has demonstrated generally effective instructional and intervention techniques for writing (e.g., see Wendling & Mather, 2009 for a summary), little is known about gender-specific writing intervention techniques. The most informed interventions will, of course, be contingent on an understanding of why these differences arise in the first place. Nonetheless, future intervention research should investigate possible gender-by-instruction interactions that may be used to maximize the writing skills for both males and females. Until then, because of the limited research, it is most likely advantageous to make sure that males are exposed to the most effective writing instruction.

Limitations

The present findings should be interpreted in light of the limitations. First, we only investigated reading, math, and writing composites. There may be gender differences in more specific aspects of these broad achievement areas (e.g., math calculation vs. applied math problems) that are masked when scores are composited. Alternatively, it may also be that gender differences that exist at the composite level are underestimated in tests of more basic skills; such differences in basic skills may not completely reflect the large difference in writing and underestimate true educational consequences. For example, some tests of basic skills have been found to show differential achievement prediction across gender (Kranzler, Miller, & Jordan, 1999).

Second, previous research has shown female advantages in cognitive variables such as processing speed (Camarata & Woodcock, 2006). We were not able to control such variables in this study, but they may be related to writing differences.

Finally, cross-sectional data across different age levels allowed us to investigate academic achievement over the course of childhood and adolescence. Longitudinal data, however, would be more suitable to fully understand effects related to development by controlling intraindividual variability. These studies are most important with regard to the gender-by-age moderation detected in this study.

Conclusions

The most striking finding of the present study was the female advantage in writing observed in a representative sample of typically developing children and adolescents across all ages. A steady developmental trend was detected; the female advantage increases with age. The National Commission on Writing in America's Schools and Colleges (2003) has recognized writing as the "Neglected R." In the process of this neglect, an important finding has emerged: Males, who comprise half of the student population, are at a relatively large disadvantage in one of the most important skills required for success in society.

References

Aiken, L. S., & West, S. G. (1991). Multiple regression: Testing and interpreting interactions. Thousand Oaks, CA: Sage.

Arbuckle, J. L. (1995–2011). AMOS 20.0 user's guide. Crawfordville, FL: AMOS Development Corporation.

Berninger, V. W., & Fuller, F. (1992). Gender differences in orthographic, verbal, and compositional fluency: Implications for assessing writing disabilities in primary grade children. Journal of School Psychology, 30, 363–382.

Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Gender differences in severity of writing and reading disabilities. Journal of School Psychology, 46, 151–172.

Camarata, S., & Woodcock, R. (2006). Sex differences in processing speed: Developmental effects in males and females. Intelligence, 34, 231–252.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed.). Mahwah, NJ: Erlbaum.

DeLa Paz, S., & Graham, S. (2002). Explicitly teaching strategies, skills, and knowledge: Writing instruction in middle school classrooms. Journal of Educational Psychology, 94, 687–698.

Floyd, R. G., McGrew, K. S., & Evans, J. J. (2008). The relative contributions of the Cattell-Horn-Carroll cognitive abilities in explaining writing achievement during childhood and adolescence. Psychology in the Schools, 45, 132–144.

Gallagher, A. M., DeLisi, R., Holst, P. C., McGillicuddy-De Lisi, A. V., Morely, M., & Cahalan, C. (2000). Gender differences in advanced mathematical problem solving. Journal of Experimental Child Psychology, 75, 165–190.

Goiran-Bevelhimer, A. F. (2008). Males and writing: Implications for creating school writing curriculum and instruction that is male-friendly. Journal of Liberal Arts and Science, 13, 73–92.

Graham, S., & Harris, K. R. (1997). Self-regulation and writing: Where do we go from here? Contemporary Educational Psychology, 22, 102–114.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. Science, 269, 41–45.

Hidi, S., Berndorff, D., & Ainley, M. (2002). Children's argument writing, interest and self-efficacy: An intervention study. Learning and Instruction, 12, 429–446.

Hyde, J. S. (2005). The gender similarities hypothesis. American Psychologist, 60, 581–592.

Hyde, J. S., Fennema, E., & Lamon, S. (1990). Gender differences in mathematics performance: A meta-analysis. Psychological Bulletin, 107, 139–155.

Kaufman, A. S., Johnson, C. K., & Liu, X. (2008). A CHC theory-based analysis of age differences on cognitive abilities and academic skills at ages 22 to 90 years. Journal of Psychoeducational Assessment, 26, 350–381.

Kaufman, A. S., & Kaufman, N. L. (2005). Kaufman test of educational achievement – second edition (KTEA-II): Brief form. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., Kaufman, J. C., Liu, X., & Johnson, C. K. (2009). How do educational attainment and gender relate to fluid intelligence, crystallized intelligence, and academic skills at ages 22–90 years? Archives of Clinical Neuropsychology, 24, 153–163.

Keith, T. Z. (2006). Multiple regression and beyond. Boston, MA: Pearson Education.

King, K., & Gurian, M. (2006). Teaching to the minds of boys. Educational Leadership, 64, 56–58.

Kranzler, J. H., Miller, M. D., & Jordan, L. (1999). An examination of racial/ethnic and gender bias on curriculum-based measurement of reading. School Psychology Quarterly, 14, 327–342.

Lachance, J. A., & Mazzocco, M. M. (2006). A longitudinal analysis of sex differences in math and spatial skills in primary school age children. Learning and Individual Differences, 16, 195–216.

Lubinski, D., & Humphreys, L. G. (1990). Assessing spurious "moderator effects": Illustrated substantively with the hypothesized ("synergistic") relation between spatial and mathematical ability. Psychological Bulletin, 107, 385–393.

Lykken, D. T. (1968). Statistical significance in psychological research. Psychological Bulletin, 70, 151–159.

Malecki, C. K., & Jewell, J. (2003). Developmental, gender, and practical considerations in scoring curriculum based measurement writing probes. Psychology in the Schools, 40, 379–390.

Martin, D. J., & Hoover, H. D. (1987). Sex differences in educational achievement: A longitudinal study. Journal of Early Adolescence, 7, 65–83.

McGraw, R., Lubienski, S. T., & Strutchens, M. E. (2006). A closer look at gender in NAEP mathematics achievement and affect data: Intersections with achievement, race/ethnicity, and socioeconomic status. Journal for Research in Mathematics Education, 37, 129–150.

Mullis, I. V., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). PIRLS 2006 international report. Boston: IEA.

National Commission on Writing in America's Schools and Colleges. (2003; 2004). Neglected R: The need for a writing revolution. Princeton, NJ: College Entrance Examination Board.

Pajares, F., Valiante, G., & Cheong, Y. F. (2007). Writing self-efficacy and its relation to gender, writing motivation, and writing competence: A developmental perspective. In G. Rijlaarsdam (Series ed.) and P. Boscolo & S. Hidi (Volume eds.), Studies in writing, volume 19, writing and motivation (pp. 141–159). Oxford: Elsevier.

Persky, H., Daane, M., & Jin, Y. (2003). The nation's report card: Writing. Washington, DC: U.S. Department of Education.

Salahu-Din, D., Persky, H., & Miller, J. (2008). The nation's report card: Writing. Washington, DC: U.S. Department of Education.

Schneider, W., & Näslund, C. 1999. Impact of early phonological processing skills on reading and spelling in school: Evidence from the Munich longitudinal study. In F. E. Weinert & W. Schneider (Eds.), Individual development from 3 to 12, findings from the Munich Longitudinal Study (pp. 126–147). Cambridge: Cambridge University Press.

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Constable, R. T., Skudlarski, P., Fulbright, R. K., . . . Gore, J. C. (1995). Sex differences in the functional organization of the brain for language. Nature, 373, 607–609.

Thorndike, R. M. (2005). Measurement and evaluation in psychology and education (7th ed.). Upper Saddle River, NJ: Pearson Education.

Urbina, S. (2014) Essentials of psychological measurement (2nd ed.). Hoboken, NJ: Wiley.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2011a). The nation's report card: Writing 2011. Retrieved from http://nces.ed.gov/nationsreportcard/pdf/main2011/2012470.pdf

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2011b). Digest of education statistics. Retrieved from http://nces.ed.gov/programs/digest/d11/tables/dt11_126.asp

Wendling, B. J., & Mather, N. (2009). Essentials of evidence-based academic interventions. Hoboken, NJ: John Wiley & Sons. Associated Press.