How do Educational Attainment and Gender Relate to Fluid Intelligence, Crystallized Intelligence, and Academic Skills at Ages 22–90 Years?

Alan S. Kaufman a, James C. Kaufman b,*, Xin Liu c, Cheryl K. Johnson c

a Yale University School of Medicine, New Haven, CT, USA
b Learning Research Institute, California State University at San Bernardino, San Bernardino, CA, USA
c Pearson Assessments, Bloomington, MN, USA

Accepted 19 December 2008

Abstract

Educational attainment and gender differences on fluid intelligence (Gf), crystallized intelligence (Gc), and academic skills in reading, math, and writing were analyzed for stratified adult samples ranging in age from 22 to 90 years. The data sources were the adult portions of the standardization samples of the second editions of Kaufman Brief Intelligence Test (N = 570) and the Kaufman Test of Educational Achievement–Brief Form (N = 555). Five univariate analysis of covariance were conducted with age as the covariate. Correlational analysis supplemented the covariate analyses to better understand the relationship of the five variables to education. All variables related significantly and substantially to years of formal schooling, an important finding in view of the key nature of this background variable for conducting neuropsychological assessments, as elaborated by Heaton and his colleagues. Surprisingly, Gf related just as strongly to education as did the school-related Gc. Among academic skill areas, math correlated higher with years of formal schooling than did either reading or writing. Women significantly outperformed men on the writing test and the reverse was true for the math test; other gender differences were not significant. These analyses fill a gap in the literature regarding the nature of gender and education differences in academic skills for heterogeneous samples of normal adults between young adulthood and old age and have practical implications for neuropsychological assessment.

Keywords: Intelligence; Gender; Educational attainment; Fluid intelligence; Crystallized intelligence

Introduction

The aim of this article was to examine the relationship of two essential background variables—educational attainment and gender—to cognitive ability and academic skills in normal samples of adults across the life span. From a practical standpoint, these data are valuable for clinical neuropsychological assessment by providing the kinds of demographic data analyses that facilitate the estimation of the pre-morbid functioning of an individual with brain damage (Heaton, Taylor, & Manly, 2003; Lezak, 1995). As Heaton and coworkers (2003) note: “Ideally such estimates should take into consideration any aspects of normal people that are known to be significantly associated with (i.e., to predict) test performance” (p. 184). Heaton and colleagues (2003) indicate that educational attainment, gender, and ethnicity are the three key background variables that predict test performance and are important to investigate to enhance diagnosis of neuropsychiatric disorders. For this study, we opted to analyze two of these three variables; we did not investigate ethnic differences because that variable is multi-faceted and complex and often raises more questions than it answers (see e.g., Weiss et al., 2006).

The adult portions of the standardization samples (ages 22–90 years) of the Kaufman Brief Intelligence Test—Second Edition (KBIT-2; Kaufman & Kaufman, 2004) and the Brief Form of the Kaufman Test of Educational Achievement—
Second Edition (KTEA-II; Kaufman & Kaufman, 2005) provided the data sources to evaluate the relationships of Fluid Intelligence (Gf), Crystallized Intelligence (Gc), and academic skills to both educational attainment and gender. KBIT-2 Nonverbal and Verbal provided good measures of Gf and Gc, respectively, from the Cattell–Horn–Carroll (CHC) theory (McGrew, 2005) and the KTEA-II Brief form provided measurement of the academic skills of math, reading, and writing. These samples were chosen because they are stratified according to US Census data on important background variables and include substantial numbers of normal adults at each age group across the broad age range from young adulthood to old age. The best way to understand the relationships of background variables to adult abilities and academic skills is to use reliable, valid measures of tests administered to a large, stratified sample of adults at all ages. Such normal samples permit generalization of the results to a far greater extent than do clinical samples, specialized samples (e.g., college students), or samples of individuals from narrow age ranges (e.g., the elderly).

**Educational Attainment**

Educational attainment and IQ, especially Wechsler’s IQs, relate substantially, offering important implications for psychometric interpretation of test scores (Reynolds, Chastain, Kaufman, & McLean, 1987), neuropsychological assessment (Heaton et al., 2003; Taylor & Heaton, 2001), clinical assessment (Kaufman & Lichtenberger, 2006; Matarazzo, 1972; Wechsler, 1958), and the evaluation of elderly individuals (Ryan, Paolo, & Findley, 1991). For example, Wechsler (1958, p. 251) reported noteworthy correlations between years of schooling and WAIS Verbal (.73) and Performance (.61) scores, results replicated on the WAIS-R and WAIS-III (Kaufman & Lichtenberger, 2006, pp. 115–120; Manly, Heaton, & Taylor, 2000).

V-IQ has consistently related more strongly to educational attainment than P-IQ (Kaufman & Lichtenberger, 2006). Based on old conceptions of Gc and Gf, which defined Wechsler’s V-IQ as Gc and P-IQ as Gf (Horn, 1989; Matarazzo, 1972), these consistent results would lead to the easy interpretation that Gc has a stronger association with educational attainment than does Gf. However, as Horn (1989; Horn & Blankson, 2005) expanded and refined Gf–Gc theory, it became clear that V-IQ corresponds fairly well to Gc, but that P-IQ is an amalgam of Gf, Processing Speed (Gs), and especially, Visual Processing (Gv).

Indeed, when “purer” measures of Gf (such as matrices tasks) have been compared with measures of Gc as correlates of educational attainment, some evidence suggests that the coefficients may be almost identical in magnitude (Kaufman & Wang, 1992), although other studies disagreed (Kaufman et al., 1995; Manly et al., 2000). Consequently, one aim of this study was to compare the relationship of fairly pure measures of Gc and Gf to education in representative samples of adults between the ages of 22 and 90 years. Such data are needed to provide better estimates of pre-morbid functioning when the neuropsychological assessment battery includes tests that yield fairly pure measures of Gc and Gf (such as the Woodcock-Johnson III; Woodcock, McGrew, & Mather, 2001) rather than Wechsler’s V-IQ and P-IQ.

There are a plethora of studies with adults that have related performance on traditional cognitive tests (verbal—nonverbal or Gc–Gf) to educational attainment. In contrast, investigations are sparse that have related adults’ performance on traditional academic skills tests to years of formal schooling, especially for representative samples across the adult life span. Math has been studied more often than reading and writing. Wechsler Arithmetic correlated .44–.50 with years of formal schooling, about midway between the values for V-IQ and P-IQ (Kaufman & Lichtenberger, 2006, Table 4.6; Manly et al., 2000; Reynolds et al., 1987). Math also related substantially to educational attainment in the analysis of data from Iowa’s statewide adult program (Iowa Department of Education, 1998), with the most educated group (13+ years of schooling) scoring about .5 SD higher than those with 10–12 years of formal education and 1 SD higher than the least educated sample (seven to eight years of schooling). A second large-scale investigation of math and educational attainment (N = 1,146, age range = 15–93) yielded similar results (Kaufman et al., 1995).

The studies of Iowa Department of Education and Kaufman and coworkers (1995) also examined reading with similar results. In the Iowa study, the group with 13 or more years of schooling again scored 1 SD higher than those with a 7th–8th grade education, although high school graduates performed almost equally to the 13+ sample. In the other large-scale study, a measure of functional-adaptive reading (e.g., understanding abbreviations such as “non-smkrs”) correlated .52 with educational attainment, about the same as the coefficient for math. Also regarding educational attainment and reading, Franks (1998) found that 20 older adults with college experience (M = 68.3 years) performed significantly better than 15 older adults with no post-secondary education (M = 68.2 years) in the ability read and comprehend prose passages. In contrast to math and reading, the relationship between written expression and educational attainment in normal adult samples seems largely unexplored.

**Gender**

Gender differences have been investigated in adolescent and adult samples to facilitate the theoretical, societal, and developmental understanding of men versus women cognitive functioning (e.g., Hyde, 2005) and, like studies of educational
attainment, to aid practitioners in the clinical, neuropsychological, and psychometric interpretation of an adult’s IQ test profile (Heaton et al., 2003; Kaufman & Lichtenberger, 2006, pp. 96–101). Similar to the educational attainment research, the most compelling data on gender differences in normal adults between young adulthood and old age have come from Wechsler’s adult scales, with men performing substantially better than women on measures of spatial ability (Gv), such as Block Design, and women thoroughly outperforming men on measures of processing speed (Gs), like Digit Symbol (Manly et al., 2000; Reynolds et al., 1987). Men also have usually performed significantly better than women on Wechsler’s Arithmetic subtest for samples of adolescents and adults (Kaufman & Lichtenberger, 2006, pp. 98–99).

To be sure, gender differences in cognitive ability have been investigated with a wide variety of group-administered and individually administered tasks spanning a diversity of ages (see e.g., meta-analyses by Hyde, 1981, 2005). However, virtually the only studies that have systematically examined gender differences in cognitive ability for carefully stratified samples of normal adults across the adult life span are investigations of the standardization samples of tests such as the WAIS and its revisions, or of other individual tests developed for adolescent and adult populations (Kaufman, Chen, & Kaufman, 1995; Kaufman & Wang, 1992).

The best way to understand the research on Gf–Gc and gender in adults is to examine the literature that specifically investigated valid measures of Gf and Gc, although these studies of adults are usually limited to college students or samples of gifted individuals (e.g., Hyde, 2005). An overview of these numerous studies indicates that differences in Gc (defined as verbal ability) are typically quite small for children and adults (Hedges & Nowell, 1995), although when significant differences do occur they tend to favor women (Hyde, 1981, 2005). On measures of Gf (defined as performance on Raven’s matrices tests), a meta-analysis of 57 studies revealed a clear-cut developmental trend: There were no meaningful gender differences at ages 6–14 years, but men performed significantly higher than women (.33 SD) at ages 15 through old age (Lynn & Irwing, 2004). In a previous study of Gf and Gc that used the first edition of the K-BIT (Kaufman & Kaufman, 1990) at ages 4–90 years, gender differences were trivial for Gf and Gc.

Higher performance on math tests by men than women in adult samples (not samples of school-age children) is a well-established research finding. Willingham and Cole (1997) found that men earned higher scores by .36 SD units for the SAT Mathematics college admissions test, .25 SD units for the ACT Mathematics college admissions test, and .63 SD units on the GRE Quantitative graduate school admissions test. These findings occur even within the field of mathematics; there is a .87 SD unit advantage for men on the GRE Mathematics subject test (only given to those applying for graduate school in a mathematics-related discipline). The explanations and magnitude of this gender difference are varied and often controversial (e.g., Gallagher & Kaufman, 2005; Hyde, 1981, Table 2; Hyde, 2005; Hyde, Fennema, & Lamon, 1990). Some researchers (e.g., Gallagher & De Lisi, 1994; Gallagher & Kaufman, 2005; Davis & Carr, 2002), for example, point to men using more unconventional solutions, as opposed to solutions that were taught in school. Such “short cuts” may affect test performance but would likely not affect class performance—and, indeed, women score the same or better than men in mathematics class work (Bridgeman & Wendler, 1991; Kessel & Linn, 1996).

Hyde and coworkers (1990), in a meta-analysis of 56 studies of complex math problem-solving, found an interesting developmental relationship: At ages 5–14 years there were essentially no gender differences; at ages 15–18 men scored .29 SD higher; and at ages 19–25 men scored .32 SD higher. This developmental trend mirrors the trend identified for measures of Gf (Lynn & Irwing, 2004). It is also consistent with developmental findings for Wechsler’s Arithmetic subtest, which typically does not yield differences for children on the WISC or its revisions (e.g., Jensen & Reynolds, 1983), but produces substantial differences in favor of men on Wechsler’s adult scales at ages 16 through old age (Kaufman & Lichtenberger, 2006). Similarly, adolescent and adult males scored significantly higher than females (.25 SD) on the measure of math in Kaufman and co-worker’s (1995) study, and also in the Iowa state-wide study, but the effect size was small (.12 SD) (Iowa Department of Education, 1998).

In contrast to the numerous investigations of gender differences on math, there are far fewer studies of adult gender differences in reading. Generally, gender differences in reading are small or non-existent (Hyde, 2005, Table 1; Iowa Department of Education, 1998; Kaufman et al., 1995). Male–female comparisons in written expression and related abilities indicate that females generally outperform males on tests of spelling and language in high school (Feingold, 1988) and on essay tests used to select college and graduate students (e.g., Breland & Griswold, 1982; Bridgeman & Bonner, 1994; Bridgeman & McHale, 1996). Some studies, however, have reported no significant differences between adult men and women on measures of writing, such as an investigation of the SAT-II Writing Subject Test in which there were no significant gender differences on either the multiple choice or essay portions of the test (Breland, Kubota, & Bonner, 1999). In Bromley’s (1991) investigation of written production across the 20–86 year age range, men and women did not differ in most aspects of written production (e.g., word output, sentence complexity), but women performed significantly better than men on readability across the broad age range.
Overall, studies of gender differences in academic skills for adults are based on investigations of specialized samples (e.g., college students) or have used experimental measures of reading or writing ability, but have not been based on normal, heterogeneous populations of adults across the life span and have not relied on traditional, well-normed, construct-valid measures of word recognition, reading comprehension, math computation, math applications, or written expression. (Bromley, 1991, who included a normal sample of men and women of ages 20–86 years, is a notable exception.)

---

**Research Questions**

1. What is the relationship of Gf and Gc to educational attainment and gender for a stratified sample of adults, ages 22–90, covarying chronological age?

2. What is the relationship of math, reading, and writing to educational attainment and gender for a stratified sample of adults ages 22–90, covarying chronological age?

---

**Materials and Methods**

**Participants**

Two samples of adults, each ranging from young adulthood to old age, constituted the participants for this study. The samples comprise the portions of the standardization samples for the KBIT-2 (N = 570) and the KTEA-II Brief Form (N = 555) that ranged in age from 22 to 90 years. Of the 570 adults in the KBIT-2 sample, 362 were also included in the KTEA-II sample for an overlap of 63.5%. This overlap included only two individuals at the youngest age group (22–25), with the number of common cases ranging from 38 to 53 for ages 26–35 through 81–90.

Both samples for this study matched the 2001 U.S. Census data on the variables of gender, geographic region, ethnicity, and socioeconomic status (years of schooling). Participants were divided into the following nine age groups (sample sizes are shown for KBIT-2 sample before the slash and for the KTEA-II sample after the slash: 22–25 [n = 50/115], 26–35 [n = 100/60], 36–45 [n = 100/60], 46–55 [n = 60/60], 56–65 [n = 60/60], 66–70 [n = 50/50], 71–75 [n = 50/50], 76–80 [n = 50/50].
The KBIT-2 adult sample included 291 women (51.1%) and 279 men (48.9%) and comprised 65 African Americans (11.4%), 52 Latinos/as (9.1%), 434 European Americans (76.1%), 10 Asian Americans (1.8%), 8 Native Americans (1.4%), and 1 person from an “Other” ethnicity (0.2%). Overall, for the KBIT-2 sample 16.3% completed 8–11 years of schooling, 31.8% graduated high school, 25.1% had 1–3 years of college, and 26.8% were college graduates; no one completed fewer than 8 years of formal education. By age, only 7%–10% of the samples ages 22–55 completed less than 12 years of schooling compared with 30%–32% of the older samples (ages 71–90). (Education data are from Kaufman & Kaufman, 2004, Tables 4.3 and 4.4.)

The KTEA-II adult sample included 282 women (50.8%) and 273 men (49.2%) and comprised 70 African Americans (12.6%), 50 Latinos/as (9.0%), 417 European Americans (75.1%), 8 Asian Americans (1.4%), 9 Native Americans (1.6%), and 1 person from an “Other” ethnicity (0.2%). Table 1 provides the same educational attainment distributions for the KTEA-II that were provided for the KBIT-2. Overall for the KTEA-II sample, 18.7% completed 8–11 years of schooling, 29.9% graduated high school, 29.0% had 1–3 years of college, and 22.3% were college graduates; no one completed fewer than 8 years of formal education. As with the KBIT-2 sample, relatively few individuals aged 22–55 failed to graduate high school (10%–12%) compared with 28%–38% of those older than 70. (Education data are from Kaufman & Kaufman, 2005, Tables 5.10 and 5.14.)

These age differences occurred because participants in each age group were selected to match U.S. Census data on educational attainment, so the education differences from age to age reflect real societal differences. It was, therefore, important to control for age when examining the relationship of gender and educational attainment to Gf, Gc, and academic skills.

As with the WAIS-III sample (The Psychological Corporation, 1997), both adult normative samples were limited to “normal” individuals. Adults were excluded from the sample if any one of the following was true: Stroke or transient ischaemic attack, seizure disorder, alcoholism, multiple sclerosis, Parkinson’s, dementia, schizophrenia, or other psychosis, chemotherapy, treatment for depression, unconscious for 5 min or more.

**Instruments**

**KBIT-2**

All data presented here for the KBIT-2 measures of Gc and Gf are from the test manual.

**Verbal (Gc)**

The Verbal score is based on performance on two subtests: Verbal Knowledge and Riddles. Verbal Knowledge is a 60-item measure of receptive vocabulary and range of general information about the world; Riddles is a 48-item measure of verbal comprehension, reasoning, and vocabulary knowledge. Split-half reliability coefficients for adults ranged from .88 to .94 (mean = .93). Test–retest reliability (four-week interval) was .89 for ages 22–59 (n = 70) and .92 for ages 60–89 (n = 68). The Verbal score correlated .81 with V-IQ and .79 with FS-IQ on the WAIS-III for 67 adults aged 20–48 years.

**Non-verbal (Gf)**

The Non-verbal score is a 46-item Matrices subtest composed of several item types that involve relationships among both meaningful and abstract stimuli. Easy items require the person to choose which picture best completes a visual analogy (e.g., a carrot goes with a rabbit just as a bone goes with a dog). Most of the 2 × 2 and 3 × 3 Matrices items use abstract stimuli that demand non-verbal reasoning and flexibility in applying a problem-solving strategy. Split-half reliability coefficients for adults ranged from .87 to .93 (mean = .91). Test–retest reliability was .85 to .89. WAIS-III correlations were .79 (P-IQ) and .83 (FS-IQ).

**KTEA-II Brief**

All KTEA-II Brief data presented here are from the test manual.

**Reading**

The Reading score is based on performance on two parts: Recognition and Comprehension. The 37 Recognition items require reading and pronouncing irregular words that do not strictly follow phonetic rules, ensuring that word recognition, or reading vocabulary, is measured more so than decoding ability. Most of the 46 Comprehension items require reading a passage and giving oral answers to literal or inferential questions. Some items require response to commands given in printed statements (e.g., “Turn your head”). Split-half reliability coefficients for adults ranged from .91 to .99 (mean = .95). Test–retest reliability (four-week interval) was .93. Reading correlated .88 with Reading Composite on the KTEA-II Comprehensive Form at ages 19–25 years (n = 297).
Math

The Math subtest consists of 67 items that cover a broad range of computation and application skill categories. The person writes solutions to mathematical problems printed in a response booklet and responds orally to items that focus on real-life application of mathematical principles. Split-half reliability coefficients for adults ranged from .92 to .96 (mean = .94). Test–retest reliability (four-week interval) was .91 to .92. Math score correlated .87 with Math Composite on the KTEA-II Comprehensive Form.

Writing

The Writing subtest is an interactive task presented within the context of writing and editing a newsletter. The 46 items assess communication of ideas in writing without requiring reading and with minimal influence of reasoning or creativity. Skills measured include adding punctuation and capitalization, filling in missing words, completing sentences, writing original sentences, and spelling regular and irregular words. Split-half reliability coefficients for adults ranged from .87 to .91 (mean = .91). Test–retest reliability (four-week interval) was .75 to .86. Writing correlated .87 with Written Language Composite on the KTEA-II Comprehensive Form.

Data Analysis

Because the mean standard score on the KBIT-2 and KTEA-II Brief form is set at 100 for each normative age group, it was necessary to use a common yardstick, or “reference” norms group, for everyone in the adult samples to permit comparisons across age. For both tests, this reference group was selected as ages 46–55, the approximate midpoint of the wide age range of the adult samples. Raw scores earned by all individuals aged 22–90 years on the two KBIT-2 subtests and on the three KTEA-II Brief subtests were converted to standard scores (mean = 100, SD = 15) using the pertinent norms for ages 46–55.

Five univariate analysis of covariances (ANCOVAs) were conducted using the two cognitive measures (Gf, Gc) with N = 570 as dependent variables and the three measures of academic skill (math, reading, writing) with N = 555. Adjustments were made for the covariate of age (the nine age groups between 22–25 and 81–90). Independent variables were gender and educational attainment (8–11 years of formal schooling, high school graduate or general educational development (GED), some college, college graduate). The design provides main effects of gender and education, and an interaction effect of gender and education, after adjustment for age. Alpha level was set at .01 to control for the error introduced by conducting five analyses simultaneously.

The assumptions of ANCOVA are homogeneous regressions across groups. SPSS multivariate analysis of variance was used to test whether the slopes of age on each dependent variable (DV) were homogeneous across different gender or education groups. Specifically, the test for homogeneity of regression was conducted via the test for the interaction between IV and age with IVs and age adjusted for. Using an alpha level of .01, all interactions of gender X age and education by age were non-significant, showing no violation of homogeneity of regression in any of the ANCOVAs.

Also, Pearson product–moment correlation coefficients were computed between educational attainment and the five DVs; education was converted from a categorical variable to a continuous one that denotes years of schooling (i.e., 8, 9, 10, 11, 12, 14, 16). The values 8–12 correspond to the exact number of years completed in school; 14 denotes 1–3 years of college; and 16 denotes college graduate or greater. Coefficients were computed between educational attainment and: (a) Gf and Gc for the KBIT-2 sample of 570 adults, and (b) math, reading, and writing for the KTEA-II sample of 555 adults. Correlations were also computed for different age groups composed of at least 110 subjects (ages 22–35, 36–55, 56–70, and 71–90). However, only the correlations for the total samples were compared to determine whether the coefficients with educational attainment differed significantly. Hotelling’s (1940) t for dependent samples was computed, first, to compare the relationship of Gc vs. Gf to educational attainment, and next, to make pair-wise comparisons of academic skills. In the latter analyses, the Bonferroni correction was applied, using a family-wise alpha level of .05.

Results

The results of the five univariate ANCOVAs are summarized in Table 1. The covariate of age is important for Gf ($\eta^2 = .201$), but not for Gc, consistent with research results that have found Gf to be vulnerable across the life span, whereas Gc is maintained into old age (Horn, 1989; Kaufman & Lichtenberger, 2006). Similarly, age was an important covariate for math and writing but not reading, as age accounted for about 8% of the variance for math and writing, but less than 1% of the variance in reading. (For an in-depth study of aging differences on Gf, Gc, and the academic skill variables for the present samples, see Kaufman, Johnson, & Liu, 2008.)
The main effect of education was significant for both Gc and Gf, covarying age; they had almost the identical moderate effect size, with education accounting for about 18% of the variance in both Gc and Gf. Education was also a significant main effect for all the three academic skills, with each DV showing a moderate effect size (accounting for 21%–22% of the variance in reading and writing and a substantial 36.8% in math). The main effect of gender was significant for math (men scored higher) and writing (women scored higher), but not reading, Gf, or Gc. Although math and writing were significant at \( p < .001 \), both effect sizes were small. None of the gender \( \times \) education interactions reached significance at \( p < .01 \). Table 2 presents the age-adjusted mean standard scores for all DVs by gender. Men scored 4.1 points (0.27 SD) higher than women on math, whereas women outscored men by 8.8 points (0.59 SD) on writing.

The actual and age-adjusted mean standard scores on Gc and Gf are presented in Table 3 for the four educational groups. The means increased with increasing education, as expected, but contrary to prediction, the relationship was not greater for Gc than Gf. Differences between all possible pairs of education groups in Gc and Gf were tested using polynomial contrasts (\( df = 1 \), 561). Five of the six contrasts were significant at \( p < .001 \) level for both Gc and Gf; the only contrast not to reach significance at the .001 level was “some college” vs. “college graduate,” although the more educated group did score significantly higher on Gf (\( p < .01 \)). Table 4 shows the actual and age-adjusted mean standard scores on academic skills for the educational groups. The means increase with increasing education, as expected, with the extreme education categories of 8–11 years of schooling vs. college graduate differing in their age-adjusted means by 25.7 points on math (1.7 SD), 22.5 points on writing (1.5 SD), and 20.9 points on reading (1.3 SD). Differences between all possible pairs of education groups were tested using polynomial contrasts (\( df = 1 \), 546). All six contrasts were significant at \( p < .001 \) level for math and reading; for writing, five of the six were significant at \( p < .001 \) and “high school graduate” vs. “some college” was significant at \( p < .01 \).

### Table 2. Actual and age-adjusted standard score means for women and men on fluid intelligence (Gf), crystallized intelligence (Gc), and academic skills

<table>
<thead>
<tr>
<th>CHC ability</th>
<th>Female</th>
<th>Male</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gc</td>
<td>N</td>
<td>Mean</td>
<td>Age-adjusted</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Actual</td>
<td></td>
</tr>
<tr>
<td>Gc</td>
<td>291</td>
<td>95.0 (14.2)</td>
<td>94.3</td>
</tr>
<tr>
<td>Gf</td>
<td>291</td>
<td>95.0 (17.1)</td>
<td>94.3</td>
</tr>
<tr>
<td>Math</td>
<td>282</td>
<td>97.2 (13.8)</td>
<td>97.4</td>
</tr>
<tr>
<td>Writing</td>
<td>282</td>
<td>101.8 (15.6)</td>
<td>101.5</td>
</tr>
<tr>
<td>Reading</td>
<td>282</td>
<td>100.1 (14.5)</td>
<td>100.1</td>
</tr>
</tbody>
</table>

**Note:** Age-adjusted mean standard scores are in bold print. Standard deviations of actual standard scores are in brackets next to the actual means. n.s. = non-significant.

### Table 3. Actual and age-adjusted standard scores on crystallized intelligence (Gc) and fluid intelligence (Gf) for different educational levels

<table>
<thead>
<tr>
<th>Educational attainment</th>
<th>N</th>
<th>Mean Gc</th>
<th>Actual</th>
<th>Adjusted</th>
<th>Mean Gf</th>
<th>Actual</th>
<th>Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>8th–11th grade</td>
<td>93</td>
<td>83.6 (14.0)</td>
<td>84.0</td>
<td>79.2 (17.3)</td>
<td>83.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school graduate</td>
<td>181</td>
<td>92.3 (12.2)</td>
<td>92.1</td>
<td>92.7 (16.4)</td>
<td>92.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some college</td>
<td>143</td>
<td>99.2 (12.9)</td>
<td>99.0</td>
<td>99.8 (16.1)</td>
<td>98.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-year degree</td>
<td>153</td>
<td>103.0 (14.1)</td>
<td>102.3</td>
<td>106.2 (14.4)</td>
<td>104.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Age-adjusted mean standard scores are in bold print. Standard deviations of actual standard scores are in brackets next to the actual means.

### Table 4. Actual and age-adjusted mean standard scores on math, writing, and reading for different educational levels

<table>
<thead>
<tr>
<th>Educational attainment</th>
<th>N</th>
<th>Mean Math</th>
<th>Actual</th>
<th>Adjusted</th>
<th>Mean Writing</th>
<th>Actual</th>
<th>Adjusted</th>
<th>Mean Reading</th>
<th>Actual</th>
<th>Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>8th–11th grade</td>
<td>104</td>
<td>85.2 (12.4)</td>
<td>86.7</td>
<td>83.3 (17.8)</td>
<td>85.6</td>
<td>88.6 (16.9)</td>
<td>89.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school graduate</td>
<td>166</td>
<td>95.1 (11.6)</td>
<td>95.8</td>
<td>96.6 (13.7)</td>
<td>94.9</td>
<td>97.3 (11.7)</td>
<td>97.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some college</td>
<td>161</td>
<td>104.6 (11.7)</td>
<td>103.2</td>
<td>100.7 (15.2)</td>
<td>99.8</td>
<td>103.4 (11.6)</td>
<td>103.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-year degree</td>
<td>124</td>
<td>113.4 (11.6)</td>
<td>112.4</td>
<td>108.10 (14.5)</td>
<td>108.1</td>
<td>110.1 (11.4)</td>
<td>110.0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Age-adjusted mean standard scores are in bold print. Standard deviations of actual standard scores are in brackets next to the actual means.
Table 5 presents the correlational analysis between educational attainment, as a continuous variable, and standard scores on the five variables. Coefficients ranged from .34 to .69 across five age groups; all values were significant at $p < .001$. Math correlated highest with education (.63 for the total sample), with the other variables correlating in the mid- to high-.40s. Correlations were highest at ages 56–70 (mean $r = .54$) and lowest at ages 22–35 (mean $r = .41$). Based on Hotelling’s $t$ for dependent samples, the correlations for Gf and Gc for the total sample did not differ significantly [$t(567) = 1.205$]. However, the coefficient for math was significantly greater than the values for reading [$t(552) = 5.137$] and writing [$t(552) = 5.282$].

**Discussion**

Although much data have accumulated on gender and educational differences in adult samples, most of the investigations to date have been on homogeneous groups, such as college students or applicants to graduate schools, on limited samples of “older” and “younger” adults, or on available (often small) adult samples included in experimental research. When large-scale studies have been conducted with stratified samples of adults across a wide age range, the variables have typically been Wechsler’s V-IQ and P-IQ (which provide a good measure of Gc but a poor measure of Gf) or non-traditional tests of academic skill such as functional-adaptive measures of reading and math (Kaufman et al., 1995). Whereas the study of the original K-BIT (Kaufman & Wang, 1992) provided Gf-Gc gender and education relationships, literature on traditional measures of academic skills for normal, heterogeneous adult samples has been sparse. The present study helps fill that gap.

**Gender Differences**

The most important findings in this study pertain to adults’ writing ability, especially the significant gender differences (Table 2). Previous investigations of reading and math have been conducted with normal samples of adults across the life span (Iowa Department of Education, 1998; Kaufman et al., 1995), but no one has investigated measures of written expression with normal adults spanning young adulthood to old age. This study used a reliable, valid, well-normed, interactive measure of writing that does not require reading ability or much reasoning and that measures the mechanics of writing (punctuation and capitalization, grammar, spelling, and the ability to write original sentences). Women performed better than men (.59 SD using age-adjusted means) on this writing task, an effect size that is larger than the effect size for math (.28 SD favoring men). The writing difference favoring women is consistent with previous research on writing, typically with homogeneous samples such as college students, but the effect size, although small, is still considerably larger than has been found previously. For example, women outscored men by .12 SD within each of three ethnic groups (White, African American, Latino) on the Graduate Management Admissions Test (GMAT) Writing Assessment (Bridgeman & McHale, 1996).

The better performance by men than women on math is both consistent with past research on adults’ math performance and similar in magnitude. The effect size of .28 SD is virtually identical to the values reported by Hyde and coworkers (1990) in their meta-analysis of complex math problem-solving (.29 SD at ages 15–18 and .32 SD at ages 19–25), and by Kaufman and coworkers (1995) in their investigation study of functional-adaptive math abilities (.25 SD). The Iowa state-wide study reported differences in the same direction but the effect size was a bit smaller (.12 SD) (Iowa Department of Education, 1998). The present findings of better performance by men than women on a reliable and valid math test for a stratified normal sample of young adults through old age, therefore, reinforce previous results in the literature, most of which are based on young adult samples, often college students.

In contrast to the significant gender differences on math and writing, no significant difference between men and women emerged for reading, a measure of both word recognition and reading comprehension. That finding of no meaningful gender difference on reading is entirely consistent with previous research that assessed adults across the adult life span (Iowa Department of Education, 1998; Kaufman et al., 1995) and with the bulk of the literature on young adult samples.
usually college students (Hyde, 2005). The lack of gender differences on Gc and Gf is only partly consistent with prior investigations. The Gc finding agrees with past research (Hedges & Nowell, 1995). The lack of a Gf gender difference in Gf, although consistent with the Gf result obtained for K-BIT Matrices (Kaufman & Wang, 1992) is inconsistent with the .33 SD advantage for men on Raven’s matrices tests at ages 15 through old age reported in Lynn and Irwing’s (2004) meta-analysis.

Educational Attainment Differences

Years of formal schooling related significantly and substantially to Gf, Gc, and the three academic skills. This is a common sense finding but one, nonetheless, that has not been demonstrated very often with normal adult samples across the life span, particularly with measures of academic skills. For the present KBIT-2 adult sample, Gf and Gc related equally well to educational attainment. Their correlations with number of years of formal schooling (.44 for Gc and .48 for Gf) did not differ significantly, and the difference between the age-adjusted mean standard scores earned by the most educated sample (college graduates) and the least educated sample (high school dropouts) were comparable (Table 3): 18.3 points for Gc (1.22 SD) and 20.8 points for Gf (1.39 SD).

These results accord well with the comparison of Gf and Gc on the K-BIT, which correlated .59 and .64, respectively, with highest grade completed for ages 20–90 (Kaufman & Wang, 1992). However, the finding of about equal relationships for Gf and Gc with educational attainment is inconsistent with the Wechsler adult literature, which finds higher correlations for V-IQ (usually about .60) than P-IQ (usually about .40) (Kaufman & Lichtenberger, 2006). The difference may pertain to the feasibility that Wechsler’s P-IQ measures Gv and Gs more so than Gf, and therefore, does not provide a true Gf–Gc comparison.

Math correlated higher with educational attainment (r = .63) than did any other variable studied (Table 5), and this value is significantly greater than the correlations for reading (.49) and writing (.48). This finding is sensible because neither reading nor writing is very dependent on specific facts and concepts taught in the last few years of high school or in college; they are skills that relate significantly to education, but the basic ingredients to read efficiently and write effectively are taught long before the end of high school. In contrast, one’s math ability is always dependent on specific facts, concepts, and processes that are learned incrementally as one progresses through algebra, geometry, trigonometry, and (for some individuals) calculus. The measure of math in this study includes a number of advanced items that require the specialized knowledge taught in high school and college math courses.

Although the present analyses cannot be used to affirm causality, it is sensible to conclude that the significantly higher correlation between math and educational attainment, relative to the coefficients for reading and writing, suggests causality between educational attainment and math skills—that is, a noteworthy portion of the relationship is due to the specific influence of formal math instruction on an adult’s later math skills.

Another interesting finding concerning educational attainment and math skills concerns the age trend in Table 5: The highest coefficients are at ages 56–70 (mean r = .54) and the lowest are at ages 22–35 (mean r = .41). This finding is counterintuitive. If formal schooling has a direct, specific affect on an adult’s performance on academic skills tests, one would expect that affect to be felt more acutely shortly after the school experience (in young adulthood) and not many years afterwards. The reason for the higher correlations for older than younger age groups (and a drop in the coefficients for ages 71–90 compared with 56–70) is as elusive as the explanation for another counterintuitive, but perhaps related, research result—that the role played by genetics in affecting IQ is substantially larger for adults than for children (Plomin & Petrill, 1997); and among adults, the heritability stays high through age 70, but drops for the elderly (Bouchard, 1996; Kaufman & Lichtenberger, 2006, pp. 31–33).

Applications to Clinical Neuropsychology

The effects of education and other demographic variables on cognitive, memory, and academic test performance are well known, and have been used by Heaton and coworkers (2003) to develop demographically corrected T-scores for the WAIS-III and Wechsler Memory Scale—Third Edition (WMS-III). Education has traditionally accounted for the largest portion of variance on neuropsychological tests relative to other demographic variables, and has provided useful information regarding a person’s best estimate of pre-morbid functioning and most accurate psychodiagnostic classification (e.g., Diehr, Heaton, Miller, & Grant, 1998; Heaton, Grant, & Matthews, 1991; Heaton et al., 2003). Present data extend the results to the KBIT-2 and KTEA-II, with the age-adjusted means and standard deviations for four educational groups, presented in Tables 3 and 4, serving as demographic norms for clinicians who use these tests for neuropsychological assessment. The data on the KTEA-II are especially valuable because they extend the work of Heaton and colleagues into a new area of assessment, namely the evaluation of academic skills. The relationship between educational attainment and Gf–Gc in a large normal representative sample of adults between early adulthood and old age is, we believe, particularly noteworthy.
The data on gender differences are especially noteworthy for neuropsychological assessment. Typically, gender differences “are, for the most part, trivial” (Heaton et al., 2003, p. 196). Trivial differences were, indeed, found for the two cognitive variables and for reading, but women outperformed men in writing by more than .5 SD, and men had higher scores on math. The gender difference found in mathematics testing is quite well known (e.g., Gallagher & Kaufman, 2005), perhaps related to the affect of stereotyping on female performance (e.g., Inzlicht & Ben-Zeev, 2000; Quinn & Spencer, 2001; Spencer, Steele, & Quinn, 1999). However, the female advantage in writing is not well known. This finding suggests that gender needs to be treated as an important demographic variable when attempting to determine an adult’s pre-morbid writing ability. The means and standard deviations in Table 2 provide demographic norms that will aid in the separate determination of pre-morbid writing skills for men versus women.

The overall results for reading ability are also of interest. Reading was the only academic skill that did not relate significantly to the age covariate (see also Kaufman et al., 2008) and it did not produce gender differences. These findings are generally supportive of the occasional clinical practice of using tests of word reading to estimate pre-morbid ability (Lezak, 1995).

Conclusions

Gf and Gc related equally well to educational attainment suggesting that previous studies with Wechsler’s adult scales, which consistently yielded higher correlations for V-IQ (a good measure of Gc) than P-IQ (a poor measure of Gf), may not be directly relevant to purer measures of Gf and Gc. This finding is directly applicable to neuropsychological assessment when test batteries use tests based on CHC theory, such as the WJ III, rather than the more traditional Wechsler scales.

All measures of adult academic skills related significantly and substantially to educational attainment. However, math correlated higher with education than did reading and writing (or Gf and Gc) suggesting that at least part of the relationship between educational attainment and math skills is causal in nature. For neuropsychological assessment, a person’s years of formal schooling will provide an excellent estimate of pre-morbid math skill.

Women performed substantially better than men in writing across the life span (.59 SD). In contrast, men outperformed women in math from young adulthood to old age (.28 SD), and there was no significant gender difference in reading. These findings add considerably to the gender literature because of the stratified samples of adults at all age levels and the reliable, valid measures of writing, math, and reading that were used. The substantial gender difference among adults in writing, previously unknown, have direct implications for pre-morbid estimates of written expression for men vs. women.

Conflict of Interest

Alan S. Kaufman earns royalties from Pearson Assessments for the KBIT-2 and the KTEA-II Brief Form and is co-author of both of these tests. Xin Lin and Cheryl Johnson are both employees of Pearson Assessments, publisher of the KBIT-2 and KTEA-II Brief Form.

Acknowledgements

The authors are grateful to Mark H. Daniel, PhD, for his contributions to this paper.

References
