Heritability of Word Recognition in Middle-Aged Men Varies as a Function of Parental Education

William S. Kremen,[1,8] Kristen C. Jacobson,[2] Hong Xian,[3] Seth A. Eisen,[3,4] Brian Waterman,[5] Rosemary Toomey,[6] Michael C. Neale,[2] Ming T. Tsuang,[1,6] and Michael J. Lyons[6,7]

Although it is of lifelong importance, reading ability is studied primarily in children and adolescents. We examined variation in word recognition in 347 middle-aged male twin pairs. Overall heritability (a^2) was 0.45, and shared environmental influences (c^2) were 0.28. However, parental education moderated heritability such that a^2 was 0.21 at the lowest parental education level and 0.69 at the highest level; c^2 was 0.52 and 0.00, respectively. This constitutes a parental education × environment interaction. The higher heritability was due to a decrease in the magnitude of shared environmental factors, rather than an increase in the magnitude of genetic factors. Other cognitive studies have reported gene × environment interactions, but patterns may differ as a function of age or specific cognitive abilities. Our results suggest that shared environmental factors in families with low parental education have long-lasting effects on word recognition ability, well beyond any critical period for developing reading proficiency.

KEY WORDS: Heritability; middle-aged; moderator; reading ability; twin study; word recognition parental education.

Given the obvious importance of developing good early reading skills, it is not surprising that there is an extensive body of literature examining genetic and environmental influences on reading ability in childhood and adolescence. Although reading is an activity with lifelong-not just childhood-significance, little is known about how genetic and environmental influences affect individual differences in reading ability in midlife, well past any critical period for developing proficiency. Parental education or social class can also moderate genetic and environmental influences on individual differences in particular cognitive abilities during childhood and adolescence (Fischbein, 1980; Rowe *et al.*, 1999; Scarr-Salapatek, 1971; Turkheimer *et al.*, 2003). It is not known, however, whether parental education moderates genetic and environmental influences on components of reading ability specifically, or whether such moderator effects are manifest in midlife.

Heritability estimates for different components of reading ability from twin (Brooks *et al.*, 1990; Knopik and DeFries, 1999; Matheny and Dolan, 1974; Stevenson *et al.*, 1987; Thompson *et al.*, 1991) and adoption (Cardon *et al.*, 1990; Wadsworth *et al.*, 1995) studies in which participants were unselected for

[1] Department of Psychiatry, Institute of Behavioral Genomics, University of California, San Diego, La Jolla, CAUSA.
[2] Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA.
[3] Department of Internal Medicine, Washington University School of Medicine, St. Louis, MO.
[4] Department of Psychiatry, Washington University School of Medicine, St. Louis, MO.
[5] Waterman Research Solutions, St. Louis, MO.
[6] Harvard Institute of Psychiatric Epidemiology and Genetics, Boston, MA.
[7] Department of Psychology, Boston University, Boston, MA.
[8] To whom correspondence should be addressed at Department of Psychiatry, University of California, San Diego, 9500 Gilman Drive (MC 0603), La Jolla, CA 92093-0603. Tel: 858-822-2393. Fax: 858-822-5856. E-mail: wkremen@ucsd.edu.

reading disability range from 0 to 0.69. Reading is a complex skill requiring multiple component processes, and the measurement of different component processes is one likely reason for the variability in results. Assessment measures have included: word recognition; individual components underlying recognition such as phonological decoding, orthographic coding, and phoneme awareness; comprehension; and composite indices. In addition, little is known about genetic and environmental influences on these processes in middle adulthood because most studies include participants ranging from 7 to 20 years or age. The measure used in the present study was the reading subtest of the Wide Range Achievement Test-3 (WRAT; Wilkinson, 1993); this achievement subtest is a measure of single word recognition accuracy. (We use the acronym WRAT to refer specifically to the reading subtest.) Heritability estimates for word recognition based on best fitting biometrical genetic models from previous studies with child and adolescent samples unselected for reading difficulties or samples selected as controls for participants with reading difficulties range from 0.19 to 0.49 (Brooks *et al.*, 1990; Cardon *et al.*, 1990; Knopik and DeFries, 1999; Stevenson *et al.*, 1987; Wadsworth *et al.*, 1999; Wadsworth *et al.*, 1995).

One study by Gayán and Olson (2003) reported substantially higher heritability of 0.85 for a combined sample of controls and participants with reading difficulties, although they noted that values for the two subsamples did not differ. Contrary to most other research, there was evidence for significant non-additive genetic influences and no evidence for shared environmental effects in the latter study. However, there were few twins from families with very low socioeconomic status (SES) in their sample. The authors acknowledged that both the presence of non-additive genetic influences and the use of a relatively high SES sample could result in overestimates of heritability. In addition, their estimates for the heritability of word recognition were based on multivariate analysis of 12 different cognitive variables, and they included a timed recognition test as part of their latent word recognition measure. The speed factor reflects another process that could have increased the heritability estimate of the latent word recognition trait, given that heritability for the timed test was 0.90. Finally, multivariate measurement models such as the one used by Gayán and Olson (2003) are useful because they allow variation due to test error to be separated from nonshared environmental influences on the latent trait. However, these

measurement models also provide higher heritability estimates than those obtained from univariate studies because the total variance of the latent trait is less than that of the measured variables (Kendler *et al.*, 1993). Thus, the proportion of genetic variance on the latent trait (i.e, the heritability) is higher than that for the measured variable.

The different, yet complimentary theoretical perspectives of Scarr (1992) and Bronfenbrenner and Ceci (1994) suggest that it may also be important to take parental education into account when estimating genetic and environmental influences on individual differences in midlife word recognition ability. Scarr theorized that the impact of family environments would be greater if one were below a certain threshold in the range of environmental quality. Thus, heritability would be likely to decrease and shared environmental effects would be likely to increase for individuals experiencing less favorable family environments. Essentially, the reasoning behind this notion was that above a certain threshold, the family environment is adequate enough and will have little influence on individual differences in IQ or other cognitive abilities ("good-enough environments" hypothesis).

In thinking of development as, in part, a series of transactions between the individual and the environment, Bronfenbrenner and Ceci (1994) reasoned that the full expression of genetic traits was dependent on environmental influences that tend to bring out that trait. Thus, better environments should increase heritability. On the other hand, they also argued that environmental influences are not reduced with better environments. This position contrasts with the view of Scarr (1992) that environmental influences tend to be more important only at the less favorable end of the environmental continuum.

As noted by Rowe *et al.*, (1999), the position of Bronfenbrenner and Ceci (1994) may not be tenable from a quantitative genetics perspective because phenotypic variance is, by definition, the sum of genetic and environmental variance. If the former increases, the latter must decrease. If heritability increases with more enriched environments, it is possible, however, that shared environmental influences would be diminished while the impact of unique environmental influences would remain unchanged or even increase. This type of scenario could partially accommodate the view of Bronfenbrenner and Ceci. In either case, both conceptualizations suggest a gene × environmental context interaction.

Environmental context might take the form of parental involvement or the quality of the family

socioemotional climate. Children are more likely to view reading as enjoyable when there is a positive family socioemotional climate, and children whose early experiences with reading are enjoyable are more likely to become frequent readers (Baker *et al*., 1997). Reading to young children has a positive influence on the development of literacy (Scarborough and Dobrich, 1994). Parental involvement in teaching children about reading is also associated with improvement in reading skills (Leslie and Allen, 1999; Lonigan and Whitehurst, 1998). Some longitudinal data suggest that reading to children or exposure to books influences language skills directly, and these language skills influence subsequent reading ability; on the other hand, parental teaching about reading influences emerging literacy skills directly, which in turn, influence word reading (Sénéchal and LeFevre, 2002).

Scarr-Salapatek (1971) and Fischbein (1980) showed that social class—one element of which is typically parental education—moderates the heritability of IQ or other cognitive abilities in children and adolescents. Unfortunately, these studies were limited by imprecise measures of social class (Fischbein, 1980; Scarr-Salapatek, 1971), absence of zygosity measures (Scarr-Salapatek, 1971), or small sample size (Fischbein, 1980).

Rowe *et al*., (1999) examined the issue of parental education as a moderator of genetic and environmental influences on receptive vocabulary in the large and methodologically rigorous National Longitudinal Study on Adolescent Health (Add Health). They examined performance on the Peabody Picture Vocabulary Test (PPVT) in 1909 adolescent twin and sibling pairs. The PPVT is a receptive vocabulary measure that is frequently used as an estimate of verbal IQ. When the sample was considered as a whole, genetic influences accounted for 57% of the variance in PPVT scores and 13% of the variance was accounted for by shared environmental influences. As predicted, however, these influences were both moderated by level of parental education. Specifically, the heritability ranged from 0.74 in families with the most educated parents to 0.26 in families with less well-educated parents. In contrast, common environment accounted for between 0% and 23% of variance in PPVT scores across families with more versus less parental education, respectively.

Turkheimer *et al*., (2003) examined this issue with Wechsler IQ scores in 7-year-old twins from the National Collaborative Perinatal Project. Similar to the Rowe *et al*. findings, Turkheimer *et al*. found that heritability was 0.72 in high-SES twins but only

0.10 in the low-SES twins. Common environmental influences accounted for only 15% of the variance in high-SES twins, but it accounted for 58% in low-SES twins. To the extent that better educated parents provide a more enriched environmental context for the development of reading ability, these results suggest that genetic factors are more important, and shared environmental factors are less important in "better" environments.

A limitation of the Rowe *et al*., (1999) study is that the components of variance underlying the increase in heritability and the decrease in relative shared environmental influences were left unspecified. Heritability can increase if the overall magnitude of genetic factors increases across different environmental contexts and environmental variance remains constant (or increases at a lower rate relative to the increase in genetic variance), resulting in a larger proportion of variation due genetic factors. Alternatively, if the overall magnitude of environmental factors decreases across different environmental contexts, yet the magnitude of genetic factors remains constant, heritability would also increase. Note that in the first example, the overall phenotypic variance would increase whereas in the second example the overall phenotypic variance would decrease across different environmental contexts. Because Rowe *et al*., (1999) used the regression-based DeFries–Fulker method of estimating gene × environment interactions, rather than a formal structural equation approach, it is unclear whether the reported increase in heritability and decrease in shared environment is due to an absolute increase in genetic variance, an absolute decrease in shared environmental variance, or both. Using structural equation modeling, Turkheimer *et al*., (2003) did show that increased heritability of IQ in their high-SES twins was a function of both an increase in genetic variance and a decrease shared environmental variance.

The goals of the present study were to: (1) determine the extent of genetic, and shared and unique environmental influences on word recognition in midlife; (2) determine whether parental education moderates genetic and environmental influences on individual differences in midlife word recognition ability such that heritability would be higher among those whose parents were more educated; and (3) examine which genetic and environmental components of variance account for the hypothesized gene × environment interaction by means of formal structural equation modeling (Neale *et al*., 2002).

METHODS

Participants

Participants were members of the Vietnam Era Twin (VET) Registry, a nationally distributed sample of male-male twin pairs in which both members served in the military during the Vietnam era (1965–1975). Zygosity was assigned using questionnaire and blood group methods (Eisen *et al.*, 1989); this approach has been shown to achieve approximately 95% accuracy when compared to DNA analysis (Nichols and Bilbro, 1966; Peeters *et al.*, 1998). The twins responded to a zygosity questionnaire that included 20 questions covering the following categories: (l) zygosity self-assessment (1 item); (2) global similarity as children ("peas-in-a-pod" question; 1 item); (3) specific similarity as children (12 items); and (4) frequency with which during childhood the twins were confused by others (6 items). Scores on each item ranged from $+1$ (complete concordance for similarity) to -1 (concordance for dissimilarity). Zygosity was determined on the basis of logistic regression using discriminating variables from among the 20 items supplemented with blood group typing data from military records. Twins were classified as monozygotic (MZ) if: (1) the probability of being MZ based on the logistic regresssion was > 0.90; and (2) they were concordant for both blood groups (ABO and Rh). They were classified as dizygotic (DZ) if: (1) the probability of being MZ was ≤ 0.50; or (2) either blood group was discordant. Those with a probability of being MZ that was > 0.50 and ≤ 0.90 who were also blood group concordant were classified as indeterminant zygosity. A complete description of the Registry's construction is available elsewhere (Eisen *et al.*, 1987; Henderson *et al.*, 1990).

In the Harvard Drug Study, 8169 randomly selected VET Registry twins were interviewed by telephone; there were over 3300 pairs (> 6600 individuals) in which both members of a pair participated (Tsuang *et al.*, 2001). In the present study, a random selection of 348 pairs ($N = 179$ MZ, $N = 169$ dizygotic DZ) was recruited from the Harvard Drug study sample for participation in a separate twin study of vulnerability to alcoholism. The present analyses were conducted with this subsample of 348 twin pairs, as a word recognition test was given during the alcohol vulnerability study. Twins were not selected for the alcohol vulnerability study on the basis of alcohol or drug use; however, only twins without service in Vietnam were recruited for the present study to avoid the potential confounding influence of combat exposure. To be included, both members of a pair had to agree to participate. These participants were flown in from around the country for a day-long series of assessments in Sacramento, CA or Boston, MA. Participants were given their choice of study site. After a complete description of the study to participants, written informed consent was obtained at the study sites. There were 176 complete MZ pairs and 169 complete DZ pairs who participated; 181 pairs were tested in Boston, 163 in Sacramento, and 1 pair in their hometown. In virtually all cases, both members of a pair came together to the same site. In addition, 3 MZ twins participated without the participation of their co-twin.

As a gauge of the representativeness of the sample examined in the present study, we provide some demographic data for the entire portion of the Harvard Drug study sample that did not have service in Vietnam ($N = 4066$) and for the twins in the present analyses. We did not perform statistical tests because the present sample is a subset of the Harvard study sample. Except for one variable, these data are based on information at the time of the Harvard study (completed in 1992); the exception is the twins' education which was collected in another study in 1987. In each case, the first number shown represents the larger Harvard study sample and the second number represents the sample for the present study. Note that income is based on a 1–10 scale and that education measures were coded in a slightly different manner than in the present study: age (41.4 [SD = 2.9], 40.8 [SD = 3.1]); education (13.6 [SD = 2.1], 13.9 [SD = 2.1]); parental education (11.3 [SD = 2.5], 11.5 [SD = 2.3]); married (77, 78%) employed (96, 98%); income (7.0 [SD = 2.4], 7.2 [SD = 2.30]). Based on these data, the present sample does appear to be representative of our larger VET Registry sample.

We had valid word recognition scores from 692 individual twins. This included complete data from both twins in 176 MZ pairs and 168 DZ pairs, and data from only one twin (with missing data from his co-twin) in 4 families (3 MZ, 1 DZ). However, there was 1 MZ pair that had missing data on parental education and was excluded from the analyses. Thus, our final sample consisted of data from 347 twin pairs (343 complete; 4 with data from 1 twin) for a total of 690 individual twins. The mean age of these participants was 47.9 years (SD = 3.3; range = 41–58); 92.2% were non-Hispanic white, 5.5% were African-American, 1.9% were Hispanic, and 0 4% were other; 96.7% were high school graduates 32.8% were college graduates; 92.4% were employed full-time and 1.6% were employed part-time.

Measures

Word recognition ability was assessed with the reading subtest of the WRAT-3 (Wilkinson, 1993). The WRAT is a widely used and well-validated instrument. Version 3 has a normative database of nearly 5000 people. Participants must read and correctly pronounce a series of increasingly difficult words. Words are read out of context so that there are no cues as to meaning or pronunciation. The version utilized in the present study has a median internal consistency of 0.91 based on coefficient αs across all age groups, and a stability coefficient (corrected for attenuation) of 0.98 after an average of 37 days (Wilkinson, 1993). WRAT standard scores have a mean of 100 and SD of 15 in the general population. In our sample of 690 individuals, mean WRAT = 97.3 (SD = 10.6) with a range of 51–118. Although the mean of our sample is close to the population mean, the variability is somewhat reduced. This is likely to be a function of military enrollment characteristics. Military recruits of this era took the Armed Forces Qualification Test, and those below the 10th percentile were excluded from service. In addition, higher SES individuals were substantially under-represented (Boulanger, 1981). Consequently, the sample tends to generate what is essentially a truncated normal distribution of cognitive abilities.

Parental education was based on each twin's report of the highest level of formal education completed by his mother and father (possible range = 0–20 years). Correlations of the within-pair report of mother's, father's, and average parental education were 0.84–0.86 (p_s < 0.0001), indicating substantial agreement across twins. We, therefore, averaged the reports of mother and father education across twins. In addition, mother's and father's education were moderately correlated ($r = 0.49$; $p < 0.0001$). Consequently, our measure of parental education for these analyses was the average of both mother's and father's educational attainment, using the averaged reports from both twins. Twin reports on both mother and father education, from both members of the twin pair, were available from 299 pairs (86.7% of the sample). In six additional pairs (1.7%), data were available for mother's education alone. For the remaining 11.6% ($N = 38$ pairs plus 4 unpaired twins), data on both mother's and father's education were available from at least one twin.

In this sample, father's mean level education was 11.1 years (SD = 3.1), mother's mean level education was 11.6 years (SD = 2.6), and the mean average parental education level was 11.3 years (SD = 2.7). For ease of interpretation in our structural equation models, the average parental education level variable that was used as our moderator effect was re-coded to range from 0 (representing the lowest level of average parental education, i.e., 0 years in this sample) to 1 (representing the highest level of average parental education in the sample, i.e., 19 years in this sample).

Statistical Analysis

The standard twin ("ACE") model estimates the proportion of phenotypic variance due to additive genetic effects, (a^2), common familial or shared environmental effects (c^2), and unique or unshared environmental effects (e^2) (Neale and Cardon, 1992). Note that a^2 refers to the heritability, also commonly indicated by h^2. The c^2 term reflects environmental influences common to twin pairs, whereas e^2 reflects individual-specific environmental influences, plus measurement error.

Parameters were estimated using the maximum-likelihood-based structural equation modeling program, Mx (Neale et al., 1999). An advantage of using Mx is that it allows the user to fit complex models to raw data, rather than to covariance matrices of MZ and DZ twins. This feature allows for the inclusion of measured variables that can be used as continuous moderators of genetic and environmental influences (Purcell, 2002). To examine whether the magnitude of genetic and environmental influences on word recognition varies across level of parental education, we compared the fit of three different models:

(1) The *no-change-in-variance ACE model* assumes that variance in word recognition remains stable across different levels of parental education.
(2) The *scalar ACE model* assumes that the total variance in word recognition changes across different levels of parental education *and* that the changes in genetic, shared environmental, and nonshared environmental influences are proportional. Thus, estimates of standardized a^2, c^2, and e^2 remain stable at all levels in this model.
(3) The *moderated ACE model* allows genetic and environmental influences to change differentially across levels of parental education, and therefore predicts different estimates of a^2, c^2, and e^2.

The overall fit of the each model is calculated by comparing twice the negative log-likelihood (−2LL) from the relevant model to the −2LL of a model that fits the data perfectly (a saturated model). The difference between these models is distributed as χ^2. Comparisons between two competing nested models can be done using the Likelihood-Ratio χ^2 Test (LRT) and Akaike's Information Criterion (AIC; Akaike, 1987). The LRT is calculated as the difference in −2LL between two models, and is distributed as a χ^2 (Neale and Cardon, 1992). The AIC, which is an index of both goodness-of-fit and parsimony, is calculated by taking the difference in −2LL between the saturated model and the model that is being tested, and then calculating that difference minus 2 times the difference in degrees of freedom (df) between the two models. Lowest (largest negative) AIC values indicate a better balance of goodness-of-fit and parsimony.

Figure 1 presents the moderated ACE model for the univariate case. In this figure, the rectangles refer to the measured twin 1 and twin 2 scores on WRAT reading. Circles represent the latent influences of genetic (A), shared environmental (C) and nonshared environmental (E) effects on variation in WRAT scores, and the moderated genetic, shared, and nonshared environmental influences (A′, C′, C′, and E′, respectively). Mean WRAT scores are represented by triangles. The measured moderator variable (level of parental education) is represented by diamonds. Estimated parameters are represented by a, c, e, a', c', e', m (for means), and mm (for parental-education-moderated means). In this model, genetic influences are correlated 1.0 for MZ twins and 0.5 for DZ twins. Shared environmental influences, by definition, are correlated 1.0 across twins, regardless of zygosity. Unique (nonshared) environmental influences (which include measurement error) are, by definition, uncorrelated across twins.

As can be seen in Figure 1, the moderated parameters (a', c', and e') are literally multiplied by the moderator variable (parental education) for each twin. Likewise, the model also allows for mean level WRAT scores to be moderated by level of parental education. In the no-change-in-variance model, the moderated a', c', and e' parameters are set to zero, and the model simply estimates a, c, and e in the standard fashion. For the scalar and moderated ACE models, the moderated a', c', and e' parameters are estimated. In the scalar model, a non-linear con-
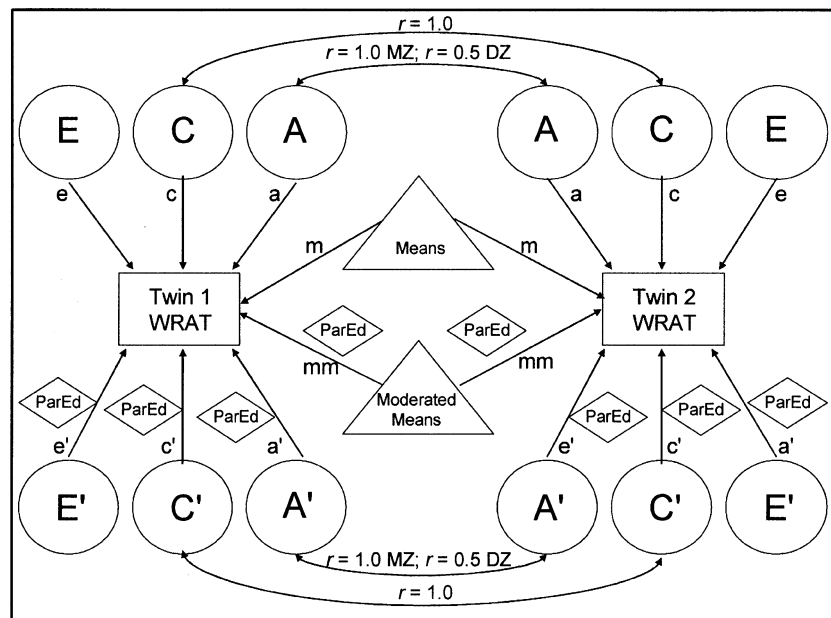

Fig. 1. Moderated ACE Model for the Univariate Case. Rectangles = measured twin 1 and twin 2 WRAT reading scores. Circles = latent influences of genetic (A), shared environmental (C) and nonshared environmental (E) effects on variation in WRAT scores, and moderated genetic (A′), shared environmental (C′), and nonshared environmental influences (E′). Triangles = mean WRAT scores. Diamonds = measured moderator variable (level of parental education). Estimated parameters are represented by a, c, e, a', c', e', m (for means), and mm (for parental education moderated means).

straint is placed on the total covariance matrix (which is based on both a, c, e, and a', c', and e' parameters), making estimates of the standardized a^2, c^2, and e^2 constant at all levels of the moderator. Thus, this model estimates only an overall change in variance, and can be compared to the no-change-in-variance model using a LRT with 1 df. Because there is no constraint placed on the covariance matrix of the moderated ACE model, the fit of that model can be compared to the fit of the scalar model using the LRT with 2 df. Given the possible confounding main effects of parental education, it is also important that all three models can include moderated means (i.e., mean WRAT increases or decreases with increased level of parental education). The best-fitting model is considered to be the model with the fewest parameters that does not manifest a statistically significant decline in its ability to fit the data.

Prior to investigating the effects of parental education level as a moderator of genetic and environmental influences on word recognition, we performed preliminary analyses to test whether parental education moderated mean-level WRAT scores. We also performed preliminary analyses to test for the significance of non-additive or dominance genetic effects (d^2) using an "ADE" model. A limitation of the classic twin design is that it is unable to simultaneously estimate the effects of shared environmental influences (C) and non-additive genetic influences (D). Therefore, we compared the fit of the ACE model to the ADE model, and examined the significance of the C and D influences. Unfortunately, these two models are not nested, and therefore cannot be compared directly.

Table I. Descriptive Statistics and Correlations in MZ and DZ Twins

Variable	MZ ($n = 353$)		DZ ($n = 337$)		
	M	SD	M	SD	p-value[a]
Age (years)	47.8	3.6	47.9	2.9	0.75
Twin Education (years)	14.2	2.2	14.1	2.1	0.75
Parental Education (years)	11.5	2.6	11.1	2.7	0.19
WRAT (standard score)	96.9	11.1	97.7	10.1	0.41

Note: MZ = Monozygotic; DZ = Dizygotic; WRAT = Wide Range Achievement Test-Version 3 reading subtest.
[a]Denotes p-values adjusted for twin-pair clustering (i.e., for the fact that not all observations are independent).

RESULTS

Descriptive Statistics

Table I shows relevant sample characteristics for MZ and DZ twins, as well as the p-values from tests examining differences in mean levels across zygosity (significance levels for tests of zygosity differences in twin's level of education and WRAT scores were adjusted for the correlation of standard errors using generalized estimating equations). None of these variables differed between MZ and DZ twins. WRAT scores were more highly correlated in MZ ($r = 0.77$; 95% CI = 0.70–0.82) than in DZ ($r = 0.51$; 95% CI = 0.39–0.62) twins, suggesting a genetic influence on word recognition. There were only modest within-twin correlations for WRAT scores and parental education, and these correlations were similar for MZ ($r = 0.21$, $p < 0.003$) and DZ ($r = 0.27$, $p < 0.0001$) twins. Significance levels for these latter correlations were adjusted for twin clustering to account for non-independence of observations.

For descriptive purposes, we also divided the sample into approximate thirds based on the average education level of the parents. The three categories were: < 11 years of education (34.3% of twin pairs); 11–12 years of education (33.7%); and >12 years of education (32.0%). Table II shows the mean WRAT scores and twin correlations for each of these three subgroups. As expected, mean WRAT scores increased with higher levels of education. Variance of WRAT scores also decreased linearly across groups. DZ twin correlations were lower in the high education group, but no consistent linear pattern appeared among the MZ twin correlations. Using the standard method of calculating heritability based on $2(r_{MZ} - r_{DZ})$ (Eaves, 1982), the heritabilities in the 3 education groups were: 0.36 (low); 0.62 (medium); and 0.56 (high). Similarly, estimates of c^2 were: 0.36 (low); 0.22 (medium); and 0.12 (high). Thus, these results suggest a pattern of decreasing variation in WRAT with higher levels of education, and a general increase in heritability and a decrease in shared environmental influence across groups. However, it should be noted that these groupings are for descriptive purposes only, and the statistics presented might differ with different methods of dividing the sample. One advantage of using Mx for the structural equation modeling is that it allows us to treat the parental education variable as a continuous measure, which gives us a more precise index of significant changes in overall variance, as well as potential changes in estimates of a^2 and c^2.

Table II. Descriptive Statistics for Sample Divided into 3 Groups based on Average Parental Education

	Low education (<11 years)	Medium education (11–12 years)	High education (≥12 years)
Individual twins (*N*)	236	233	221
MZ pairs (*N*)	55	57	63
DZ pairs (*N*)	62	59	47
WRAT (Mean, *SD*)	94.6 (11.6)	97.5 (10.6)	99.9 (8.7)
MZ twin correlation (95% CI)	0.72 (0.56; 0.82)	0.84 (0.75; 0.90)	0.68 (0.53; 0.79)
DZ twin correlation (95% CI)	0.54 (0.33; 0.69)	0.53 (0.32; 0.69)	0.40 (0.13; 0.61)

Note: MZ = Monozygotic twins; DZ = Dizygotic twins. WRAT = Wide Range Achievement Test-Version 3 reading subtest.

Structural Equation Modeling Analyses: Preliminary Analyses

Adding moderated means for MZ and DZ twins significantly improved model fit (LRT = 24.18, df = 1, $p < 0.001$), with an average increase of 18.4 WRAT points from lowest to highest parental education level. Thus, we allowed for moderated mean-level WRAT scores in all subsequent models. Although (as noted previously) we could not directly compare the standard ACE model (Model 0 in Table III) to the standard ADE model, the fit of the ADE model was worse than that of the ACE model (ACE Model: −2LL = 4996.80, df = 685, AIC = 21.8; ADE Model: −2LL = 5001.50, df = 685, AIC = −17.1). Moreover, when non-additive genetic effects were allowed in the model, they were estimated at zero. In contrast, when shared environmental influences were allowed in the model they accounted for 28% (95% CI = 3–48%) of the variance in WRAT scores and were significant. Thus, we conducted all subsequent model fitting based on the ACE model, rather than the ADE model.

Model Testing: Step 1

As described in the Introduction, an increase in heritability can arise either due to an overall increase in the magnitude of genetic influence (which would result in an overall increase in phenotypic variance), or from a greater decrease in environmental influences relative to genetic influences (which would result in an overall decrease in phenotypic variance). Thus, the first step of analysis was to test whether the overall phenotypic variance in word recognition changes across the level of parental education, and if so, whether changes reflect variance increases or decreases (see Table III). In Table III, the LRT compares the submodels to the comparison model (model 0, the standard ACE model, which assumes

that there is no change in overall phenotypic variance) to test whether adding a parameter to represent the change in variance significantly improves the model fit. Model 2 – the scalar model with proportional decreasing variance – is the best-fitting model in Table III because it is the only model with a significant LRT and it has the lowest AIC value. This scalar model indicates that genetic influences account for nearly half of the variance in word recognition ($a^2 = 0.46$; 95% CI = 0.23–0.72), whereas shared and unique environmental influences each account for a little over one-fourth of the remaining variance ($c^2 = 0.27$; 95% CI = 0.01; 0.47; $e^2 = 0.28$; 95% CI = 0.22–035). For illustrative purposes, the last two columns in Table III show the variance at the lowest (0 years) and highest (19 years) levels of average parental education. Although estimates of a^2, c^2, and e^2 are nearly identical for the standard (un-moderated) ACE model (Model 0) and the scalar decreasing variance model (Model 2), the values in these columns illustrate the difference between the two models. These columns clearly indicate that the best fitting model (Model 2) is one in which parental education moderates the variance in word recognition such that the variance decreases in participants whose parents have higher levels of education. The variation in WRAT scores among twins whose parents are least educated is much greater (180.3, 95% CI = 136.4–226.7) than variation in WRAT scores among twins whose parents are most educated (54.1; 95% CI = 31.2; 83.3). Note that because this model assumes that changes in genetic and environmental influences are proportional, the standardized estimates of a^2 and c^2 are identical at all levels of parental education. Figure 2 depicts the changes in genetic, shared, and nonshared environmental variance of WRAT scores across level of parental education based on the results from the decreasing scalar variance model.

Table III. Model Fitting for Estimating Changes in Genetic and Environmental Variance Components in Reading Ability as a Function of Parental Education: Step 1. Scalar Models[a]

Model		Fit Statistics					Variance component estimates			Variance	
#	Name	−2 LL	df	AIC[d]	LRT[e]	df	a^2	c^2	e^2	Lowest parental education[f]	Highest parental education[f]
0	No moderation[b] (95% CI)	4996.80	685	−21.8	N/A		0.45 (0.22; 0.71)	0.28 (0.03; 0.48)	0.27 (0.22; 0.34)	106.4 (94.3; 120.9)	—
1	Increase variance[c] (95% CI)	4996.80	684	−19.8	0.00	1	0.45 (0.22; 0.71)	0.28 (0.03; 0.48)	0.27 (0.21; 0.34)	106.4 (89.2;120.9)	106.4 (94.3; 123.4)
2	Decrease variance[c] (95% CI)	4984.02	684	−32.6	12.78*	1	0.46 (0.23; 0.72)	0.27 (0.01; 0.47)	0.28 (0.22; 0.35)	180.3 (136.4; 226.7)	54.1 (31.2; 83.3)

[a] All models allow for moderation of mean levels of WRAT scores (i.e., means increase with level of parental education); however, *Scalar* models assume that changes in variance components are proportional across parental education levels. [b] Model 0 represents the standard ACE model. [c] Models 1 and 2 are *Scalar* models that test whether variance in word recognition (based on WRAT scores) changes as level of parental education increases or decreases, respectively. [d] AIC = Akaike's Information Criterion. Lower (more negative) AIC values indicate *better* goodness of fit. [e] LRT = Likelihood Ratio χ^2 Test. The LRT compares models 1 or 2 with the no moderation model (model 0). Significant LRT values indicate that the model fits significantly *better* than the no moderation model. [f] Lowest and highest parental education refer to the lowest and highest levels reported in this sample; these are presented to illustrate the range of variance differences across parental education levels. There is only one value for model 0 because model 0 does not allow for changes in variance across parental education levels.

*$p < 0.0003$.

Model Testing: Step 2

Following from the step 1 results, step 2 began with the assumption that variance in word recognition decreases as level of parental education increases. We then compared the fit of different submodels to determine specifically which effects were responsible for the decrease in variance observed in step 1 (see Table IV). In other words, we tested whether the decrease in overall phenotypic variation was due to decreases in the absolute magnitude of genetic (A), shared environmental (C) or nonshared environmental (E) effects, or some combination of these three influences. The LRT statistics in Table IV reflect comparisons of these submodels to the comparison model (in this case, Model 3 in Table III, the full decreasing ACE model). In contrast to Table III these submodels have fewer parameters than the comparison model. Thus, significant χ^2 values in Table IV indicate that the submodel provides a significantly *worse* fit than the comparison model. Again, however, a lower (more negative) AIC value indicates a better model in terms of goodness-of-fit.

According to these criteria, one can see that model 5 – decreasing variance due to C and E only—is the best model in Table IV because it is the only model with a nonsignificant LRT and it has the lowest AIC value (see Figure 3 for a graphical depiction of the results from this model). It can be seen in both Table IV and Figure 3 that the overall phenotypic variance in model 5 decreases from 179.5 to 56.1 at the lowest and highest parental education levels, respectively. This result is quite similar to the decrease in variance observed in the best-fitting model from Table III (Model 2), suggesting that our moderated ACE model captured the change in variance seen in the scalar model. Table IV further shows that – consistent with our predictions – the relative magnitude of genetic factors influencing word recognition increased from a minimum of 0.21 (95% CI = 0.09–0.41) to a maximum of 0.69 (95% CI = 0.40–0.84) across level of parental education. Conversely, the relative magnitude of shared environmental influences decreased from 0.52 (95% CI = 0.27–0.68) to 0.00 (95% CI = 0.00–0.32). Technically, model 2—the decreasing variance scalar model—is the best-fitting model overall because it has the lowest AIC. However, we believe that this conclusion may be somewhat misleading and that there is good reason to consider model 5 – the decreasing CE model – as the best fitting model (see Discussion).

Table IV. Model Fitting for Estimating Changes in Genetic and Environmental Variance Components in Reading Ability as a Function of Parental Education: Step 2. Non-Proportional Models[a]

| # | Name | Fit statistics | | | | | Lowest parent education | | | Highest parent education | | | Variance | |
		-2 LL	df	AIC[b]	LRT[c]	df	a^2	c^2	e^2	a^2	c^2	e^2	Lowest parental education[d]	Highest parental education[d]
3	Decreasing ACE	4981.94	682	−30.7	N/A		0.41	0.37	0.21	0.51	0.00	0.50	207.5	37.2
	95% CIs						0.11;0.79	0.02;0.65	0.12;0.3.5	0.00;0.82	0.00;0.76	0.18;1.0	146.7;270.5	11.9;75.8
4	Decreasing AE	4986.84	683	−27.8	4.9*	1	0.79	0.00	0.21	0.51	0.00	0.49	203.7	35.3
	95% CIs						0.30;0.88	0.00;0.33	0.11;0.43	0.00;0.81	0.00;0.80	0.13;1.0	118.8;264.0	11.0;96.9
5	Decreasing CE	4983.34	683	−31.3	1.40	1	0.21	0.52	0.27	0.69	0.00	0.31	179.5	56.1
	95% CIs						0.09;0.41	0.27;0.68	0.18;0.37	0.40;0.84	0.00;0.32	0.16;0.55	137.4;225.1	35.7;82.9
6	Decreasing AC	4985.35	683	−29.3	3.41*	1	0.48	0.38	0.14	0.22	0.00	0.78	169.9	37.5
	95% CIs						0.17;0.84	0.03;0.66	0.10;0.21	0.00;0.63	0.00;0.57	0.37;1.0	146.1;256.7	24.2;76.2
7	Decreasing C	4988.63	684	−28.0	6.69**	2	0.23	0.59	0.18	0.56	0.00	0.44	163.6	67.5
	95% CIs						0.09;0.45	0.34;0.74	0.14;0.24	0.31;0.70	0.00;0.25	0.30;0.68	125.0;205.7	49.9;92.1
8	Decreasing A	4990.53	684	−26.1	8.59**	2	0.86	00	0.14	0.27	0.00	0.73	199.8	38.0
	95% CIs						0.41;0.90	0.00;0.36	0.10;0.27	0.00;0.65	0.00;0.65	0.27;1.0	114.8;248.5	22.3;100.5
9	Decreasing E	4995.03	684	−21.6	13.09***	2	0.62	0.00	0.38	0.83	0.00	0.17	121.5	90.3
	95% CIs						0.10;0.72	0.00;0.19	0.22;0.35	0.14;0.90	0.00;0.66	0.07;0.32	96.9;156.1	70.2;119.2

[a]Based on the results shown in Table III, these models start with the assumption that variance in word recognition (based on WRAT scores) decreases as level of parental education increases. Models in this table then test which variance components are responsible for the decrease in variance. All models allow for moderation of mean levels (i.e., means increase with level of parental education); however, all models are *Non-Proportional* models that allow variance components to change differentially across parental education levels.
[b]AIC = Akaike's Information Criterion. Lower (more negative) AIC values indicate *better* goodness of fit. [c]LRT = Likelihood Ratio χ^2 Test. The LRT compares the submodels to the comparison model (model 3). Submodels have fewer parameters than the comparison model. Thus, in contrast to Table III, significant LRT values indicate that the submodel fits significantly *worse* than the comparison model. [d]Lowest and highest parental education refer to the lowest and highest levels reported in this sample; these are presented to illustrate the range of variance differences across parental education levels.
*$p = 0.06$; **$p < 0.5$; ***$p < 0.002$.

DISCUSSION

In our sample of middle-aged men, when moderation effects were not considered, nearly half the variance in word recognition ability was accounted for by genetic influences. Specifically, the estimate for the heritability of word recognition was 0.45, with estimates of 0.28 and 0.27 for shared and unique environment, respectively. The present study examined the moderating effects of parental education level on both mean level of word recognition as well as on the sources of individual differences in word recognition. As expected, level of parental education was significantly associated with differences in mean-level WRAT scores. The average WRAT score among twins whose parents were least educated was 86.4, compared with an average of 104.8 among twins whose parents were most educated. Further analysis revealed a significant decrease in overall phenotypic variation across education level, and suggested that the three ACE components of variance also differed significantly as a function of level of parental education.

Based solely on the AIC and LRT values, the best fitting model was the scalar model in which there is a significant decrease in the overall phenotypic variance with higher levels of parental education, but the decreases in the A, C, and E variance components are proportional (see Figure 2). Thus, estimates of a^2, c^2, and e^2 were essentially identical at all levels of parental education. On the other hand, the decreasing CE model indicates a different pattern in which there is a substantial reduction in common environmental

influences on word recognition in individuals whose parents have higher levels of education, and it is this reduction that results in increased heritability (see Figure 3). Although this model was not statistically significantly better compared to the scalar model, our interpretation is that it nevertheless best explains the data. Supporting this conclusion is the fact that all of the models in Table IV that did not include decreasing C effects (Models 4 [AE], 8 [A] and 9 [E]) fit the data significantly more poorly than the full decreasing ACE model (Model 3), suggesting that the decrease in C is a real effect. Moreover, it can readily be seen in Figure 3 that it is the relatively dramatic decrease in the C effect that accounts for the observed increase in a^2 with increasing levels of parental education. In contrast, the magnitude of genetic variance remains stable across level of parental education (see Figure 3).

Based on the results from this model, we find evidence for higher shared environmental effects (0.52) at the lowest end of the parental education continuum compared to no shared environmental effects (0.00) at the highest end of the parental education continuum. Although the absolute magnitude of genetic influence remained stable across levels of parental education, the proportion of variation due to genetic factors (the heritability) increased from 0.21 at the lowest parental education level to 0.69 at the highest parental education level. It should be noted, however, that these differences in a^2 and c^2 are maximum differences because these estimates were
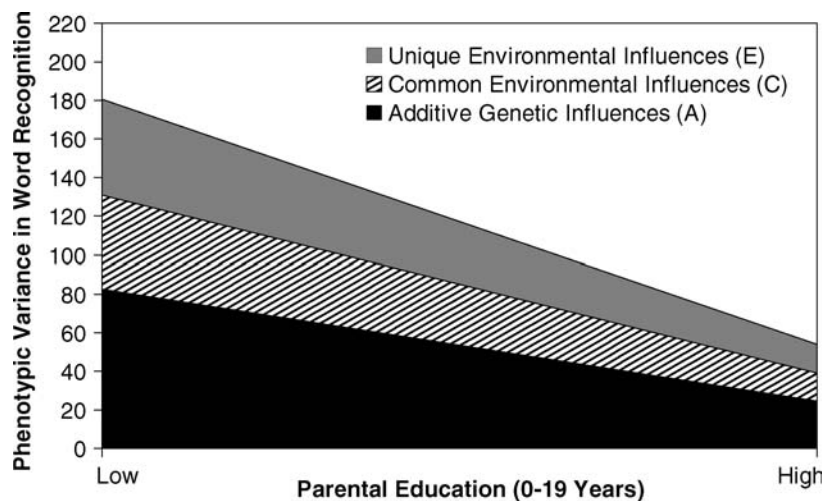

Fig. 2. Decreasing Scalar ACE Model for WRAT Reading. Assumptions are that: (1) total variance of WRAT reading scores decreases across different levels of parental education; and (2) changes in A, C, and E components of variance are proportional.
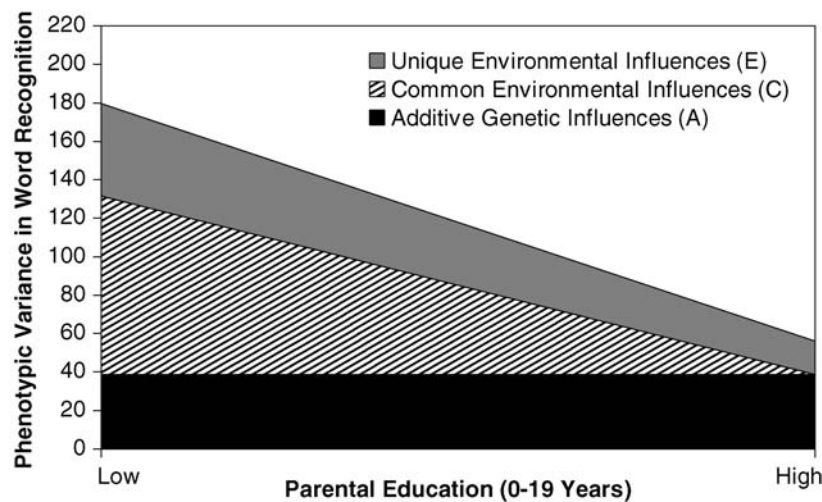
Fig. 3. Moderated (Decreasing) CE Only Model for WRAT Reading. Assumptions are that: (1) total variance of WRAT reading scores decreases across different levels of parental education; (2) there is no change in the A component of variance; and (3) changes in C and E components of variance can be non-proportional.

calculated at the two extremes of the parental education continuum.

Results from the Add Health study – which used parental education as a moderator variable in the examination of gene × environment interactions for the PPVT – further support our interpretation of the best fitting model because they strongly suggest that the failure of the decreasing CE only model to attain statistical significance over the scalar model in our study was largely a function of insufficient power rather than the lack of a meaningful interaction effect. Compared with the present study, there was a similar increase in the relative magnitude of genetic factors from 0.26 to 0.74 in the Add Health study, but actually a smaller decrease in relative magnitude of shared environmental influences (0.23 – 0.00; Rowe *et al.*, 1999). Thus, the overall interaction effect for the shared environmental influence was smaller than it was in the present study. Despite this smaller effect, re-analysis of the Add Health data by one of us (KCJ) using structural equation models implemented in Mx showed that these moderator effects were statistically significant when compared with results from a scalar variance model (results available upon request). Thus, the failure to find a significant difference between the scalar and moderated ACE models in the present study appears to be due to our significantly smaller sample size (347 twin pairs vs. 1791 twin and sibling pairs in the Add Health study) rather than a smaller interaction effect.

However, the Mx-based re-analysis of the Add Health data revealed *both* significant increasing genetic and significant decreasing environmental influences (i.e., an "X-shaped" pattern of parental education × gene and parental education × shared environment effects [results available upon request]). In the present study, we found support only for the hypothesis that the magnitude of shared environmental influences decreases with higher levels of parental education. Indeed, we formally tested this by adding an increase in genetic variance parameter to our decreasing CE model. The increase in genetic variance was estimated at zero, and the addition of this parameter did not change the model fit [LRT = 0.00, df = 1, $p > 0.99$]. Thus, the Add Health results indicate the presence of both parental education × gene and parental education × shared environment interactions, and are consistent with the theoretical formulations of both Scarr (1992) and Bronfenbrenner and Ceci (1994). Even though heritability estimates change across level of parental education in the present sample, this change is due to a change in the relative proportion of variance explained by genetic factors vis-a-vis decreasing C and E effects, and not from an actual change in the magnitude of genetic influence. Thus, the findings in the present article appear to reflect the presence of a parental education × shared environment interaction only, and are more consistent with Scarr's theory.

Paralleling the results of the Add Health study, the examination of parental SES as a moderator of childhood IQ by Turkheimer *et al.*, (2003) indicates both parental SES × gene and parental SES × shared environment interactions. What then might account for the different patterns observed in our study versus these two other studies? In addition to issues of statistical power that have been discussed, the variation in WRAT scores in the present sample (SD = 10.6) was lower than the expected population-based variance of 15.0. It is possible that this may have limited our ability to detect both significant parental education × gene and parental education × shared environment interactions. On the other hand, lack of power does not appear to be responsible for our failure to detect significant gene × shared environment interactions because the estimate for an increase in genetic variance was zero.

The between-study differences may also be a function of the different measures that were examined and/or age differences. Correlations between PPVT or Wechsler IQ scores and the WRAT are less than about 0.65, leaving a substantial proportion of unshared variance (Sattler, 1988; Wilkinson, 1993). Clearly genetic and environmental influences may operate differently on different cognitive abilities, and the moderating effects of parental education may operate differently as well. The ability of parental education to moderate genetic influence on cognitive ability may also be limited to childhood and adolescence when the effects of family environmental context on the expression of genetic potential may be more important. Specifically, the increasing variance in genetic influences at the higher end of parental education/SES in the Rowe *et al.*, (1999) and Turkheimer *et al.*, (2003) studies is consistent with the notion articulated by Bronfenbrenner and Ceci (1994) that the expression of genetic potential is more fully realized in better environments. This scenario may manifest itself as increased variance of genetic influences. At the higher end of parental education, an environment may be provided that is more favorable for the fullest development of reading skills. Conceivably, such a pattern could occur via greater exposure to intellectual stimulation, parents reading more to children, access to better schools and other resources, and even better nutrition. In our sample, however, the average age was 47.9, so twins have presumably been out of the family home for over two decades. Thus, the potential for the benefits of greater intellectual stimulation and greater resources to maximize the genetic potential of word recognition in adults might be relatively short-term.

It is interesting, however, that there was a significant effect of common environmental influences at the lower levels of parental education/SES in all three studies, and that all three found evidence in support of parental education/SES × shared environment interactions. This pattern is consistent with Scarr's (1992) notion that environmental effects have greater impact on variation in behavior and traits when the environment is less favorable. In this instance, low parental education may create a less favorable environment in regard to development of reading skills. What is particularly intriguing about the present study is that our results from adult men suggest that the disadvantageous effects of less favorable family environments may have long-lasting effects on individual differences in word recognition well into middle adulthood. This may be the case even though the twins in our sample are decades beyond any critical period for the development of reading proficiency and well beyond the time when they would have been living with their families of origin. Nevertheless, although parental education level may be considered to be an environmental factor that is particularly relevant for child and adolescent development, it is conceivable that differences in parental education are associated with differences in other salient environmental experiences throughout the life course that, in turn, affect word recognition ability in middle age. For example, individuals whose parents are less educated are less likely to go to college than individuals whose parents are more educated (Blake, 1986). Lack of exposure to a college environment may further be related to diminished enhancement of reading skills. Likewise, individuals from less educated families are likely to have less exposure to books during childhood and may thus be less likely to develop an interest in reading than individuals from more educated families; such differences may, in turn, be associated with mean level differences in reading abilities in adulthood (Scarborough and Dobrich, 1994; Baker *et al.*, 1997). Therefore, our results may additionally indicate the influence of an accumulation of later environmental experiences on the development of word recognition that are: (1) largely correlated among siblings in the same family (i.e., shared environmental influences); and (2) further related to the education level in the family of origin.

Limitations

The present study must be considered in the context of certain limitations. Our conclusions cannot necessarily be generalized to women; nevertheless, many of the findings are consistent with studies of both male and female children and adolescents and those studies did not find significant sex differences. The sample size of 347 twin pairs did not afford sufficient power to conclusively support the decreasing CE model as the best model, as described in detail above. We employed a measure of word recognition, but other aspects of reading may have somewhat different patterns of genetic and environmental influences in midlife. In addition, without the use of longitudinal and/or experimental designs, it is possible that the moderating effects we found are not due to parental education directly, but rather to an unmeasured variable that is correlated with parental education (also see below for a further discussion of sampling stratification). Nonetheless, it is of interest that we have identified a measurable environmental variable that is associated not only with differences in mean level word recognition, but is further associated with changes in the sources of individual differences in word recognition.

An additional potential limitation of these results is that our models did not incorporate the effects of assortative mating. In the present study, parents correlated 0.49 for education level, indicating some degree of assortative mating. In standard twin designs, the presence of assortative mating would predict increased DZ correlations relative to MZ correlations, which may have overstated the effects of common environment and underestimated the overall heritability in the present study (Fisher, 1918; Neale and Cardon, 1992). If there was greater assortative mating at lower levels of parental education, this could account for the pattern of decreased common environmental effects with increasing levels of parental education that was found in the present study. Unfortunately, an extended-kinship design that uses parallel data from both twins and parents is required to formally estimate the effects of assortative mating (Neale and Cardon, 1992).

Nevertheless, we were able to examine this issue to some extent by correlating the average of mother's and father's education with the absolute difference between mother's and father's education among those twin pairs who had data on both mother and father education ($N = 341$ pairs; 98.3% of the total sample). If parents who were less well-educated on average had more similar education levels (i.e., less absolute difference in education level, which would be indicated by a significant positive correlation), this might indicate the presence of differential assortative mating. The correlation between the average parental education score and the absolute difference in education level across mother and father was $r = -0.05$ ($p = 0.39$), indicating virtually no relationship between level of education and difference in education across parents. Moreover, using the sample division shown in Table II, the average absolute difference in education level between parents within each group was: 2.6 years (low education); 0.9 years (medium education); and 2.5 years (high education). Thus, parents with *either* high or low levels of education were more likely to be different from one another than parents with average levels of education. Taken together, these results suggest that differential assortative mating is not likely to be responsible for the patterns of increased C influence among twins from less educated families found in the present study.

Our formulation – as well as those of Rowe *et al.*, (1999) and Turkheimer *et al.*, (2003) – is predicated primarily on treating social class or parental education as an index of environmental context of developing twins. Clearly, however, parental education is not a pure environmental context measure. Heritability estimates for educational attainment in adults (particularly those in the age cohort of parents in the present study) generally range from approximately 0.40–0.50, with similar estimates for shared environmental effects (Fulker, 1978; Heath *et al.*, 1985; Silventoinen *et al.*, 2000; Tambs *et al.*, 1989). That being the case, we have referred to our finding as a parental education × environment interaction rather than an environment × environment interaction.

The presence of genetic influences on our environmental moderator variable (parental education) may have created a gene-environment correlation, whereby smarter (more educated) parents pass along genes related to better word recognition ability and provide more intellectually nurturing environments. In addition, the fact that parental education level is phenotypically related to WRAT scores in the whole sample ($r = 0.23$, $p < 0.001$, $N = 692$) may indicate that our results are due to an effect of sampling stratification. Specifically, there is some evidence that the heritability of some reading components processes varies across levels of general cognitive ability (e.g., Knopik *et al.*, 2002), and that components of reading ability are themselves correlated with both general

cognitive ability and parental education. Thus, it is possible that our results do not indicate the moderating effect of parental education *per se*, but rather the effects of differential heritability among individuals at different places on the word recognition continuum that is due to factors other than parental education. Unfortunately, because twins within pairs have the same value for parental education, it is impossible to test for these potential confounds. Nevertheless, it should be noted that using simulated data, Purcell (2002) estimated that the presence of gene-environment correlation had little impact on the ability to detect gene × environment interactions.

In addition, not all previous studies find effects of differential heritability across level of cognitive ability (e.g., Cherny *et al.*, 1992). Furthermore, studies of differential heritability across level of reading ability have often relied on oversampling of reading disabilities. In our sample, the average WRAT score was close to the population mean and only a small number of twins ($N = 6$) had WRAT scores below 70 (i.e., 2 SDs below the population mean), suggesting that it was not oversampled for reading disabilities. Nevertheless, it is possible that because parental education was correlated with WRAT scores in the present sample, by using parental education as a moderator variable we may have differentially selected individuals from different parts of the WRAT continuum (i.e., a sample truncation effect). However, as shown in Neale *et al.*, (1989), the effects of sample truncation are often very small when variables related to the truncation are only modestly correlated with the outcome variable (i.e., "soft selection"), as in the present study where the correlation between parental education and WRAT scores ($r = 0.23$) was modest.

It is also unlikely that a statistical artifact of sample selection accounts for the pattern of changes in C because such effects would be expected to be symmetric, such that those from the top 10% of the population would be expected to have the same altered level of C as those from the bottom 10%. In our analyses, the change in C is linear, decreasing across the range of parental education. Finally, in the present analyses level of parental education has been regressed out of the WRAT score (via the moderated means). The use of residual WRAT scores means that any shared variance between parental education and WRAT scores, be it environmental or genetic variance, has most likely been at least partially controlled. Thus, our analysis of interactions seems unlikely to be biased by any potential genetic overlap between the two measures, or by sampling stratification.

Summary

To our knowledge, the present study is the first to investigate in middle-aged twins: (1) the heritability of word recognition; and (2) the moderating effects of parental education on variation in word recognition. Consistent with previous research on other cognitive abilities in children and adolescents, we found evidence for the increased importance of shared environmental factors on individual differences in word recognition among twins from less educated families. In contrast to previous research on other cognitive abilities, we did not find evidence that the magnitude of genetic factors increased with level of parental education. Thus, whereas previous studies of children and adolescents have found both parental education/SES × gene and parental education/ SES × environment interactions, we found only the latter for word recognition in our sample of middle-aged men. The different pattern may be attributed to differences in age and/or in the particular cognitive abilities that were assessed across samples. Although the heritability of reading recognition ability during midlife was higher in individuals with better educated parents, this higher heritability was accounted for primarily by reduced variance in common environmental influences – rather than increased variance in genetic influences – with increasing parental education. This may indicate that the ability of environmental factors to moderate the expression of genetic potential on some components of reading ability is strongest in, or limited to childhood and adolescence. A corollary to this conclusion may be that, in poor environments, the impact of environmental factors on word recognition is long lasting (given the manifestation of this pattern in midlife adults). Thus, we believe our findings provide further support for the value of early environmentally-based intervention programs such as Head Start for enhancing reading and other school-related abilities in children from impoverished environments. As such, further investigation of these processes using other measures of cognitive function is warranted.

REFERENCES

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

Baker, L., Scher, D., and Mackler, K. (1997). Home and family Influences on motivations for reading. *Educational Psychologist* **32**:69–82.

Blake, J. (1986). Number of siblings, family background, and the process of educational attainment. *Soc. Biol.* **33**:5–21.

Boulanger, G. (1981). Who Goes to War?. In A. Egendorf, C. Kadushin, R. S. Laufer, G. Rothbart and L. Sloan (eds.), *Legacies of Vietnam: Comparative Adjustment of Veterans and Their Peers, Vol 4. Long-term Stress Reactions: Some Causes, Consequences, and Naturally Occurring Support Systems.* Washington, DC: U.S. Government Printing Office, pp. 494–515.

Bronfenbrenner, U., and Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental perspective: a biological model. *Psychol. Rev.* **101**:568–586.

Brooks, A., Fulker, D. W., and DeFries, J. C. (1990). Reading performance and general cognitive ability: a multivariate genetic analysis of twin data. *Person. Individ. Diff.* **11**:141–146.

Cardon, L. R., DiLalla, L. F., Plomin, R., DeFries, J. C., and Fulker, D. W. (1990). Genetic correlations between reading performance and IQ in the Colorado Adoption Project. *Intelligence* **14**:245–257.

Cherny, S. S., Cardon, L. R., Fulker, D. W., and DeFries, J. C. (1992). Differential heritability across levels of cognitive ability. *Behav. Gen.* **22**:153–162.

Eaves, L. J. (1982). The utility of twins. In V. E. Anderson, W. A. Hauser, J. K. Penry and C. F. Sing (eds.), *Genetic basis of epilepsies.* New York: Raven Press, pp. 249–276.

Eisen, S. A., Neuman, R., Goldberg, J., Rice, J., and True, W. (1989). Determining zygosity in the Vietnam Era Twin Registry: an approach using questionnaires. *Clin. Genet.* **35**:423–432.

Eisen, S. A., True, W. R., Goldberg, J., Henderson, W., and Robinette, C. D. (1987). The Vietnam Era Twin (VET) Registry: method of construction. *Acta Genet. Med. Gemellol. (Roma)*. **36**:61–66.

Fischbein, S. (1980). IQ and social class. *Intelligence* **4**:51–63.

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Translations R. Soc. Edinburgh* **52**:399–433.

Fulker, D. W. (1978). Multivariate extensions of a biometrical model of twin data. In W. E. Nance (ed.), *Progress in clinical and biological research*, Vol 24a, New York: Alan R. Liss.

Gayán, J., and Olson, R. K. (2003). Genetic and environmental influences on individual differences in printed word recognition. *J. Exp. Child Psychol.* **84**:97–123.

Heath, A. C., Berg, K., Eaves, L. J., Solaas, M. H., Corey, L. A., Sundet, J., Mangus, P., and Nance, W. E. (1985). Educational policy and the heritability of educational attainment. *Nature* **314**:734–736.

Henderson, W. G., Eisen, S. E., Goldberg, J., True, W. R., Barnes, J. E., and Vitek, M. (1990). The Vietnam Era Twin Registry: a resource for medical research. *Public Health Rep.* **105**:368–373.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., and Eaves, L. J. (1993). The lifetime history of major depression in women: reliability of diagnosis and heritability. *Arch. Gen. Psychiatry* **50**:862–870.

Knopik, V. S., and DeFries, J. C. (1999). Etiology of covariation between reading and mathematics performance: a twin study. *Twin Res.* **2**:226–234.

Knopik, V. S., Smith, S. D., Cardon, L. R., Pennington, B., Gayán, J., Olson, R. K., and DeFries, J. C. (2002). Differential genetic etiology of reading component processes as a function of IQ. *Behav. Gen.* **32**:181–198.

Leslie, L., and Allen, L. (1999). Factors that predict success in an early literacy intervention program. *Reading Research Quarterly* **34**:294–318.

Lonigan, C. J., and Whitehurst, G. J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly* **13**:262–290.

Matheny, A. P., and Dolan, A. B. (1974). A twin study of genetic influences in reading achievement. *J Learn. Disabil.* **7**:99–102.

Neale, M. C., Boker, S., Xie, G., and Maes, H. (1999). *Mx: Statistical modeling* (5th ed.). Richmond, VA: Medical College of Virginia, Department of Psychiatry.

Neale, M. C., Boker, S. M., and Maes, H. H. (2002). *Mx: Structural Modeling* (6th ed.). Richmond, VA: Virginia Commonwealth University, Department of Psychiatry.

Neale, M. C., and Cardon, L. R. (1992). *Methodology for genetic studies of twins and families.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Neale, M. C., Eaves, L. J., Kendler, K. S., and Hewitt, J. K. (1989). Bias in correlations from selected samples of relatives: the effects of soft selection. *Behav. Gen.* **19**:163–169.

Nichols, R. C., and Bilbro, W. C. J. (1966). The diagnosis of twin zygosity. *Acta Genet. State. Med.* **16**:265–275.

Peeters, H., Van Gestel, S., Vlietinck, R., Derom, C., and Derom, R. (1998). Validation of a telephone zygosity questionnaire in twins of known zygosity. *Behav. Gen.* **28**:159–163.

Purcell, S. (2002). Variance component models for gene-environment interaction in twin analyses. *Twin Res.* **5**:554–571.

Rowe, D. C., Jacobson, K. C., and Vanden Oord, E. J. C. G. (1999). Genetic and environmental influences on vocabulary IQ: parental education level as moderator. *Child Dev.* **70**:1151–1162.

Sattlier, J. M. (1988). *Assessment of children* (3rd ed.). San Diego, CA: Jerome M. Sattler.

Scarborough, H. S., and Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Dev. Rev.* **14**:245–302.

Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Dev.* **63**:1–19.

Scarr-Salapatek, S. (1971). Race, social class, and IQ. *Science* **174**:1285–1295.

Sénéchal, M., and LeFevre, J. (2002). Parental involvement in the development of children's reading skill: a five-year longitudinal study. *Child Dev.* **73**:445–460.

Silventoinen, K., Kaprio, J., and Lahelma, E. (2000). Genetic and environmental contributions to the association between body height and educational attainment: a study of adult Finnish twins. *Behav. Gen.* **30**:477–485.

Stevenson, J., Graham, P., Fredman, G., and McLoughlin, V. (1987). A twin study of genetic influences on reading a spelling ability and disability. *J. Child Psychol. Psychiatry* **28**:229–47.

Tambs, K., Sundet, J. M., Magnus, P., and Berg, K. (1989). Genetic and environmental contributions to the covariance between occupational status, educational attainment, and IQ: a study of twins. *Behav. Gen.* **19**:209–222.

Thompson, L. A., Detterman, D. K., and Plomin, R. (1991). Association between cognitive abilities and scholastic achievement: genetic overlap, but environmental differences. *Psychol. Sci.* **2**:158–165.

Tsuang, M. T., Bar, J. L., Harley, R. M., and Lyons, M. J. (2001). The Harvard twin study of substance abuse: what we have learned. *Harv. Rev. Psychiatry* **9**:267–279.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., and Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychol. Sci.* **14**:623–628.

Wadsworth, S., Fulker, D., and DeFries, J. (1999). Stability of genetic and environmental influences on reading performance at 7 and 12 years of age in the Colorado Adoption Project. *Int. J. Behav. Dev. Psychol.* **23**:319–332.

Wadsworth, S. J., DeFries, J. C., Fulker, D. W., and Plomin, R. (1995). Cognitive ability and academic achievement in the Colorado Adoption Project: A multivariate genetic analysis of parent-offspring and sibling data. *Behav. Gen.* **25**:1–15.

Wilkinson, G. S. (1993). *WRAT3: wide range achievement test manual.* Wilmington, DE: Wide Range, Inc.

Edited by Irwin Waldman