Neuropsychology
2007, Vol. 21, No. 5, 569–580

0894-4105/07/$12.00 DOI: 10.1037/0894-4105.21.5.569

Genetics of Verbal Working Memory Processes: A Twin Study of Middle-Aged Men

William S. Kremen
University of California, San Diego

Kristen C. Jacobsen
University of Chicago

Hong Xian and Seth A. Eisen
U.S. Department of Veterans Affairs, St. Louis and Washington University School of Medicine

Lindon J. Eaves
Virginia Institute for Psychiatric and Behavioral Genetics and Virginia Commonwealth University School of Medicine

Ming T. Tsuang
University of California, San Diego and Harvard Institute of Psychiatric Epidemiology and Genetics

Michael J. Lyons
Boston University and Harvard Institute of Psychiatric Epidemiology and Genetics

Genetic and environmental influences on cognitive components of reading span in 345 middle-aged male twin pairs were examined. Shared variance among word recognition (reading only), digits forward (short-term memory only), and reading span (concurrent reading plus memory) was almost entirely mediated by common genetic influences. Overall heritability was .52 for word recognition, .27 for digits forward, and .51 for reading span. All of the genetic influences on word recognition and digits forward, but only about one-half of the genetic influences on reading span, came from a common latent phenotype. The genetic influences that were specific to reading span were concluded to most likely reflect an executive function component. Implications for genetic studies of aging and prefrontal brain function are discussed.

Keywords: genetics, reading span, executive function, working memory, concurrent processing

Despite abundant evidence for substantial genetic influences on cognition, genetic factors are rarely addressed in neuropsychological studies. The purpose of this article is to examine working memory by parsing not only cognitive component processes, but also genetic and environmental components. Examination of the genetic architecture of storage and executive components of working memory may advance the study of genetic and environmental influences on executive function and on brain regions comprising prefrontal working memory circuitry.

Traditionally, parsing cognition entails examining a more complex measure after adjusting for performance on a more simple measure that involves at least one of the component processes of the former. In twin analyses, we can also account for genetic, shared environmental and unique environmental variance components (explained in Method section). If, for example, we adjust Trails B for Trails A, an implicit assumption is that the correlation between Trails A and Trails B is due to those three variance components in proportion to the size of each component for Trails A. This assumption is not necessarily valid, but genetically uninformative designs cannot address this problem.

In the present study, we examined the heritability of the Daneman–Carpenter reading span test (Daneman & Carpenter, 1980) and performed multivariate twin analyses of reading span and two other tests that tap some of the component processes involved in that effortful verbal working memory task. Reading span is a complex task in which individuals must read sentences aloud, saying one immediately after the other without any break between sentences. While reading, they must also hold in mind the last

William S. Kremen, Department of Psychiatry, Center for Behavioral Genomics, University of California, San Diego; Kristen C. Jacobsen, Department of Psychiatry, University of Chicago; Hong Xian, Department of Veterans Affairs, St. Louis, and Department of Internal Medicine, Washington University School of Medicine; Seth A. Eisen, Department of Veterans Affairs, St. Louis, and Departments of Internal Medicine and Psychiatry, Washington University School of Medicine; Lindon J. Eaves, Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University School of Medicine; Ming T. Tsuang, Department of Psychiatry, Center for Behavioral Genomics, University of California, San Diego, and Harvard Institute of Psychiatric Epidemiology and Genetics, Boston, MA; Michael J. Lyons, Department of Psychology, Boston University and Harvard Institute of Psychiatric Epidemiology and Genetics, Boston, MA.

Preparation of this article was supported in part by National Institute on Alcohol Abuse and Alcoholism Grant AA10586 and National Institute on Aging Grants AG18386, AG18384, AG22381, and AG22982. The U.S. Department of Veterans Affairs has provided financial support for the development and maintenance of the Vietnam Era Twin (VET) Registry. Numerous organizations have provided invaluable assistance in the conduct of this study, including: U.S. Department of Defense; National Personnel Records Center, National Archives and Records Administration; Internal Revenue Service; National Opinion Research Center; National Research Council, National Academy of Sciences; the Institute for Survey Research, Temple University. Most importantly, we gratefully acknowledge the continued cooperation and participation of the members of the VET Registry and their families. Without their contribution this research would not have been possible.

Correspondence concerning this article should be addressed to William S. Kremen, Department of Psychiatry, Center for Behavioral Genomics, University of California, San Diego, 9500 Gilman Drive (MC 0738), La Jolla, CA 92093-0738. E-mail: wkremen@ucsd.edu

word of each sentence. Even though there is only one performance index, reading span still constitutes a concurrent processing task (read plus remember). Word recognition (read only) and Digit Span Forward (remember only) were selected as the two other primary measures because they were included in our test battery and met the following criteria: (a) they reflect abilities that should be essential to reading span performance and (b) they are simple measures that are relatively free of other processes that are not involved in reading span. Instead of statistically adjusting reading span for the read only and remember only tests, we performed multivariate genetic analyses to determine the extent to which there were common genetic or environmental influences as well as influences that were specific to reading span.

Heritability is the extent to which genetic differences contribute to individual differences in observed behavior, that is, the proportion of phenotypic variance attributable to genetic variance. Heritabilities for digit span (forward and backward composite) have generally been moderate (.34 to .66) in middle-aged and older adults (Finkel, Pedersen, McGue, & McClearn, 1995; Hayakawa, Shimizu, Ohba, & Tomioka, 1992; Plomin, Pedersen, Lichtenstein, & McClearn, 1994). Two studies of older adults reported digits forward and backward separately; heritability for digits forward (.00 and .27) tended to be lower than for digits backward (.49 and .44; Johansson et al., 1999; Pedersen, Plomin, Nesselroade, & McClearn, 1992).

Although it is one of multiple processes involved in reading, we used word recognition to provide a simple index of reading ability in order to account for the reading component of reading span. Heritability estimates for word recognition are based primarily on child and adolescent samples; estimates range from .19 to .49 in samples not selected for reading disability (Brooks, Fulker, & DeFries, 1990; Cardon, DiLalla, Plomin, DeFries, & Fulker, 1990; Knopik & DeFries, 1999; Stevenson, Graham, Fredman, & McLoughlin, 1987; Wadsworth, DeFries, Fulker, & Plomin, 1995; Wadsworth, Fulker, & DeFries, 1999), with one outlier at .85 (Gayán & Olson, 2003). Overall heritability of word recognition was .45 in a sample that included twins in the present analyses plus a few others (Kremen et al., 2005).

Despite its relatively widespread usage, we are unaware of any reports on the heritability of the Daneman and Carpenter (1980) reading span task. One elegant study of very young adults examined the heritability of a task with some similarities to the Daneman–Carpenter task (Ando, Ono, & Wright, 2001). Using biometrical modeling, Ando et al. parsed two components of working memory: storage (word recall) and executive (sentence verification). However, the number of words recalled in their task probably cannot be said to reflect the storage function alone (e.g., as it would for digit span forward) because the task components were measured only in the context of a concurrent processing task; there was no measure of the individual components alone. To fully understand the processes involved, we think it is necessary to study the components both individually and concurrently as we do in the present study.

Bayliss, Jarrold, Gunn, and Baddeley (2003) concluded that in complex span tasks such as reading span "the ability to coordinate the processing and storage operations would be supported by the central executive component" (p. 86), consistent with Baddeley's multiple-component model (Baddeley, 1986; Baddeley & Logie, 1999). On the basis of a factor analytic study of simple and complex span tasks, Engle and Oransky (1999) suggested that the processing, that is, controlled attention, component in complex span tasks reflects Baddeley's (1986) central executive. Cowan et al. (2005) have invoked a different explanation, arguing that tests such as reading span measure the informational capacity that can be registered in one's scope of attention. They view this as a more parsimonious explanation because it requires only a single capacity rather than a combination of storage and processing components. Thus, reading span provides a true capacity measure of the focus of attention because its processing component prevents rehearsal. In contrast, a test such as digits forward makes rehearsal possible, which allows for more than one retrieval–recall cycle on a single trial (Cowan et al., 2005). On the other hand, Cowan et al. also acknowledged that tests such as reading span could reflect executive control mechanisms responsible for adjusting the scope of attention.

A long-standing literature strongly supports the notion that these kinds of executive–working memory functions are controlled by distributed prefrontal–subcortical neural circuitry (Alexander, DeLong, & Strick, 1986; Baddeley & Della Sala, 1996; Fuster, 1989; Goldman-Rakic, 1987). Bunge, Klingberg, Jacobsen, and Gabrieli (2000) examined reading span in a functional magnetic resonance imaging study by comparing three conditions: read only, remember only, and read plus remember concurrently. In their study, as well as in four earlier studies, concurrent task performance enhanced prefrontal activation relative to individual component tasks.

Understanding genetic factors underlying prefrontal–executive function and reading span performance may be important for understanding a variety of conditions. The size or density of frontal lobe regions may be more strongly influenced by genetic factors than other brain regions (Thompson et al., 2001). Also, working memory and prefrontal functions—including reading span—are impaired in several psychiatric and neurological disorders (Baddeley, Bressi, Sala, Logie, & Spinnler, 1991; Condray, Steinhauer, van Kammen, & Kasparek, 1996, 2002; Fuster, 1989; Hervey, Epstein, & Curry, 2004; Seidman et al., 1995). Finally, these results may be particularly relevant to the study of aging. Working memory is among the cognitive functions that are more susceptible to aging effects (Hasher & Zacks, 1988; Salthouse, 1990; Wingfield, Stine, Lahar, & Aberdeen, 1988), and older adults manifest deficits on standard reading span tests (May & Hasher, 1998; May, Hasher, & Kane, 1999). Prefrontal cortex also manifests the largest age-related volume reduction of any parenchymal region (Raz, 2000). Thus, the present study of middle-aged individuals has implications for understanding the genetics of cognitive and brain aging as well as cognitive deficits occurring in a variety of disorders.

Method

Participants

Study participants were drawn from 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 to Registry members by questionnaire and blood group methods (Eisen, Neuman, Goldberg, Rice, & True, 1989) that have approximately 95% accuracy compared with DNA analysis (Nichols & Bilbro, 1966;

Peeters, Van Gestel, Vlietinck, Derom, & Derom, 1998). A complete description of the Registry's construction is available elsewhere (Eisen, True, Goldberg, Henderson, & Robinette, 1987; Henderson et al., 1990).

In the now completed Harvard Drug Study, 8,169 twins were interviewed by telephone; in over 3,300 pairs (>6,600 individuals), both members of a pair participated (Tsuang, Bar, Harley, & Lyons, 2001). In the present study, 693 individuals from these 3,300+ pairs participated. They were invited to participate in a twin study of vulnerability to alcoholism, although twins were not selected on the basis of alcohol or drug use. They were randomly selected with the single caveat that only those without service in Vietnam were recruited for the present study because another study of the Registry involving only Vietnam veterans was being conducted at the same time and 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 at the University of California, Davis in Sacramento, CA and Harvard Medical School in Boston, MA. Participants were given their choice of study site. After complete description of the study to participants, written informed consent was obtained at the study sites. There were 176 monozygotic (MZ) and 169 dizygotic (DZ) pairs; 181 pairs were tested in Boston; 163 pairs, in Sacramento; and 1 pair in their hometown. In virtually all cases, both members of a pair came together to the same site. We also included data from 3 additional MZ twins whose co-twins ended up being unable to participate.

Demographic characteristics of these participants were as follows: mean age was 47.9 years (SD = 3.3; range = 41 to 58); 92.2% were non-Hispanic White, 5.5% were African American, 1.9% were Hispanic, and 0.4% were other; mean education was 14.1 years (SD = 2.2); 97% graduated high school or obtained a graduate equivalency diploma; 33% were college graduates; mean occupational level (Hollingshead, 1975) was 5.7 (SD=2.1); 98% were employed full-time; 33.5% had service or manual labor jobs, 24.4% held clerical or semiprofessional positions, and 41.1% held professional positions; 80% were married; mean income category in the mid-1990s was $60,000 to $70,000 (range: <$10,000–$100,000). There were no significant differences between MZ and DZ twins on any demographic characteristics.

In sum, participants comprised an unscreened general population sample of middle-aged men from around the country. As would be expected, given this sampling frame, their performance on basic cognitive measures was solidly average. The mean score on the reading subtest of the Wide Range Achievement Test—Version 3 (WRAT-3; Wilkinson, 1993) was 97.3 (SD = 10.6). The mean score (percentile) on the Armed Forces Qualification Test (AFQT; Grafman et al., 1988), a 100-item multiple-choice instrument administered just prior to induction into the military that constitutes an index of general cognitive ability, was 61.4 (SD = 21.3). The mean of the entire VET Registry sample was above the 50th percentile because individuals with scores below the 10th percentile were statutorily excluded from the military. The mean AFQT score for the present sample was less than one quarter of a standard deviation above the mean of the larger VET Registry sample. Mean reaction time for correct hit trials on the Conners Continuous Performance Test (CPT; Conners, 1992) was 416.6 ms (SD = 70.3), which is close to the mean for this age group based on the test's normative data. This reaction time measure provides some index of processing speed, a factor that may be associated with reading span performance (Bayliss et al., 2003).

Key Measures

Word recognition. The word recognition index of reading ability was assessed on the basis of WRAT-3 Reading standard scores.

Digit span forward. Simple storage or maintenance of information in working memory was measured with Digit Span Forward from the Wechsler Memory Scale—Revised (Wechsler, 1987).

Reading span. Reading span (Daneman & Carpenter, 1980) begins with 5 trials of 2 sentences each (2-sentence sets), and progresses to 5 trials of 5-sentence sets, and then 3 trials of 6-sentence sets. Sentences averaging just over 14 words were presented in boldface, 14-point Times New Roman font on 5.5-in. × 8-in. index cards. Participants had to read each sentence aloud. Immediately after a sentence was read, the card was turned face down and participants had to begin reading the next card. Following the last sentence in a set, a blank card was the signal for participants to recall the last word of each sentence. Before beginning the actual test, the instructions were reviewed, and a 2-sentence practice set was presented to ensure that participants understood the task. The test is discontinued if an individual does not correctly recall all of the final words on at least 3 out of 5 sets at a given set length.

Participants were asked to try to say the final words in order, but credit was given for any correct final words regardless of order. Presumably, memory traces for words in the articulatory loop will decay rapidly without rehearsal (Baddeley, 1986). Reading without pausing between sentences makes rehearsal of the last word of the previous sentence exceedingly difficult, but that would not be the case for the last word of the last sentence. Therefore, in accordance with the standard instructions, examiners strongly emphasized the importance of not saying the last word of the last sentence first.

As in some previous studies (e.g., May et al., 1999), scores were based on the total number of correct final words recalled over all trials. There were very few false positives; errors were almost entirely errors of omission.

Missing Data

All 693 individual twins had valid scores on WRAT-3 Reading, but 7 twins (4 MZ, 3 DZ) had scores greater than 2 standard deviations below the population mean (i.e., <70). We were concerned that these outliers might bias our results because these very low scores would seriously call into question a person's ability to adequately read the reading span sentences. Therefore, we excluded the WRAT-3 Reading scores for these 7 twins from our analyses, and we treated their reading span scores as missing. However, we did use their digits forward data. In addition, 7 other twins (3 MZ, 4 DZ) were missing data on either reading span and/or Digit Span Forward due to problems in administration or because the examiner judged that the participant was not putting forth adequate effort. Again, we used the non-missing data from these twins, so all 693 twins were used in the analyses. Most of the twins (n = 679; 98%) had valid data on all three measures.

Statistical Analysis[1]

Phenotypic correlations were expressed as product–moment correlations. Analyses were performed with the maximum-likelihood-based structural equation modeling program, *Mx* (Neale, Boker, Xie, & Maes, 1999). Because *Mx* allows for the fitting of models to raw data rather than to correlation matrices, individuals with partial missing data can still be included, as can individuals whose co-twins have nonvalid data. Models were compared using the likelihood-ratio chi-square (LRC) statistic and the Akaike information criterion (AIC; Akaike, 1987; Williams & Holahan, 1994). The LRC is the difference in the −2 log-likelihood (−2LL) of a comparison model and the −2LL of a nested (reduced) model. If the LRC between two models is nonsignificant, the reduced model is generally accepted as the better model. When the LRC is nonsignificant for two or more competing models, the AIC is used to determine the preferred model. The AIC indexes both goodness-of-fit and parsimony; the more negative the AIC, the better the balance between goodness-of-fit and parsimony.

We began by testing three primary trivariate genetic factor models: (a) the independent pathways model (Figure 1a); (b) the common pathways model (Figure 1b); and (c) the measurement model (Figure 1c). All three model phenotypic correlations among the measured variables and assume that additive genetic (A), dominant/epistatic genetic (D), shared or common environmental (C; environmental factors that make twins similar), and nonshared or unique environmental (E; environmental factors that make twins different) factors may each contribute to the phenotypic variation and to covariation among measures.

The independent pathways model (Figure 1a) assumes that the genes and environments influencing covariation among measures operate directly on each variable through independent genetic and environmental pathways. This model allows for the covariation between different pairs of variables to be due to different genetic or environmental influences. For example, the covariation between word recognition and reading span could be due primarily to genetic factors, whereas the covariation between digits forward and reading span could be due primarily to shared environmental factors.

The common pathways model (Figure 1b) is a nested submodel of the independent pathways model (McArdle & Goldsmith, 1990). It assumes that a single underlying latent phenotype is solely responsible for the covariation among the three measures. In this model, genetic and environmental influences on covariation are isomorphic and operate through the latent phenotype. In other words, genes and environment influence the correlation among variables via a common pathway. The lambda (λ) paths in the common pathways model correspond to factor loadings of each measure on the latent phenotype and account for the proportion of variance in each measure that is shared with the latent phenotype. A key element of the common pathways model is its prediction that the phenotypic covariance is equally apportioned into genetic and environmental components for all combinations of variables. Because one model is a submodel of the other, the common and independent pathways models can be compared by using the LRC statistic.

The measurement model (Figure 1c) is a nested submodel of the common pathways model. The critical difference between these two models is that the common pathways model allows for genetic (A_s) and shared environmental (C_s) factors that are *specific* to each measure and do not influence the covariation among measures. In contrast, the measurement model assumes that all genetic and environmental influences are operating through the latent phenotype. In this model, the only residual variance on each measure that is not explained by the latent phenotype is assumed to be measurement error and is, therefore, modeled as nonshared environmental (E_s) factors that are specific to each measure.

In all three models: (a) additive genetic factors correlate 1.0 for MZ twins and 0.5 for DZ twins (because MZ twins share 100% and DZ twins share about 50% of their genes); (b) dominant/epistatic genetic effects are correlated 1.0 across MZ twins and 0.25 across DZ twins; (c) shared environmental factors are correlated 1.0 across twins regardless of zygosity; (d) nonshared environmental factors are uncorrelated across twins; and (e) the variance of the underlying latent genetic and environmental factors is fixed at 1.0. It may not be readily apparent why DZ twins sharing 50% of their genes leads one to set $r = .50$ within DZ pairs in the models. The key is to recall that this correlation does not refer to a phenotypic correlation between traits; rather, it is the correlation of genetic factors only for the same trait within a twin pair. Essentially, additive genetic influences are correlated in direct proportion to the number of genes in common, which is .50 for DZ twin pairs. In contrast, dominance/epistasis effects are determined by a particular pairing of alleles; because they are not determined by a summation of the effect of all alleles, they are also referred to as non-additive effects. In other words, they represent an interaction between alleles on the same gene (dominance) or on different genes (epistasis). Because, on average, there is a 25% chance that DZ twins will share both alleles at a given locus, dominant/epistatic genetic influences would correlate .25.

The aforementioned correlations are not shown in the figures because the figures depict the models for only one twin in order to simplify the display. In the common pathways and measurement models, a nonlinear constraint is imposed so that the variance of the underlying latent phenotype also equals unity. Additional details about these models can be found elsewhere (Kendler, Heath, Martin, & Eaves, 1987; McArdle & Goldsmith, 1990; Neale & Cardon, 1992). Note that A, C, D, and E components are latent (not observed) variables. For example, the analysis might indicate that 40% of the variance was attributed to additive genetic influences, but it does not tell which or how many genes are involved.

Prior to testing the fit of these three models, a multivariate saturated model was fit to the raw data. This model recaptures the observed means, variances, and covariances within each zygosity group exactly. Thus, the saturated model fits the data perfectly and can be used as a comparison model to assess the overall goodness-of-fit of each of the genetic factor models. The −2LL from the saturated model was also used as the basis for calculating the AIC values for each of the genetic factor models. Finally, nested submodels of the saturated model were used to test for significant

[1] Readers who would like more detailed information on assumptions or implementation of statistical methods in behavioral genetics are referred to the introductory textbook of Plomin, DeFries, McClearn, and McGuffin (2001).

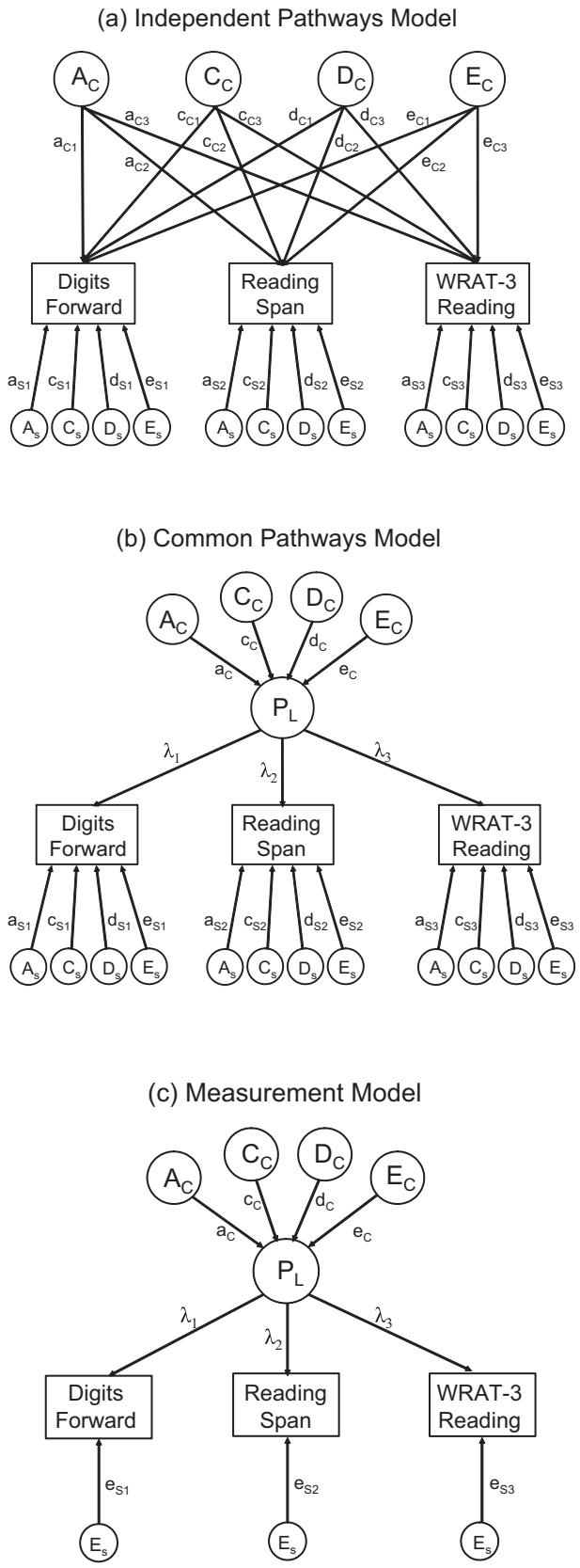

Figure 1. Trivariate genetic factor models. To simplify the display, only one twin is represented in each of Diagrams a–c. A_C = additive genetic influences; C_C = shared environmental influences; D_C = dominant/epistatic influences; and E_C = nonshared environmental influences that are *common* to all three tests. A_S = additive genetic influences; C_S = shared environmental influences; D_S = dominant/epistatic influences; and E_S = nonshared environmental influences that are *specific* to a particular test. P_L = underlying latent phenotype. Lambda (λ) paths represent factor loadings for the three measured variables on the latent phenotype. WRAT-3 = Wide Range Achievement Test—Third Edition.

differences in means and variances across Twin A and Twin B and across zygosity. Twins within pairs were randomly assigned as Twin A or Twin B.

Results

Descriptive Statistics

Descriptive statistics are presented in Table 1. To facilitate maximum likelihood estimation in *Mx*, we divided WRAT-3 Reading and reading span scores by 10 to make variances more equal across variables. Doing so affects the absolute magnitude of the −2LL, but it does not affect comparisons between models.

The genetic factors models, but not the saturated model, assume equal means and variances across twins and zygosities. MZ twins had significantly larger variance in reading span than DZ twins, $-2LL(677) = 1{,}680.81$, $LRC(1) = 7.59$, $p < .001$, but this difference was no longer significant after excluding a single outlier with a reading span score of 80, $-2LL(676) = 1{,}648.15$, $LRC(1) = 3.17$, $p = .08$. Nevertheless, we report the results on the basis of the entire sample because the overall results from the structural equation modeling analyses were the same with or without this outlier. The observed difference in variance of reading span for MZ and DZ twins will affect only the overall −2LL score of each model and will not affect comparisons between models.

Correlational Analyses

Table 2 displays the correlation matrices for MZ and DZ twins generated by the saturated model. Phenotypic (within-person) correlations ranged in magnitude from .26 to .46, indicating low-to-moderate covariation among the measures (boldface). MZ cross-twin, within-trait correlations (underlined) for digits forward and word recognition were 1.3 and 1.5 times larger, respectively, than those for DZ twins. This pattern suggests the presence of both additive genetic and shared environmental influences on these measures. In contrast, the MZ cross-twin, within-trait correlation (.55) for reading span was 6 times as large as the DZ correlation (.09), suggesting dominant or epistatic genetic influences and no shared environmental effects.

Model Fitting

Univariate genetic analyses (Table 3). A limitation of the classical twin design is that shared environmental (C) and dominant/epistatic genetic effects (D) cannot be estimated simultaneously because shared environmental influences will increase DZ

Table 1
Descriptive Statistics for Digit Span Forward, the Wide Range Achievement Test–3 (WRAT-3) Reading Subtest, and Reading Span

Variable	MZ Twin A	MZ Twin B	DZ Twin A	DZ Twin B
Digit Span Forward				
n	177	178	166	169
M	8.15	7.94	8.33	8.26
SD	2.20	2.08	2.10	2.16
WRAT-3 Reading				
n	176	175	167	168
M (divided by 10)	9.75	9.71	9.86	9.71
SD (divided by 10)	1.02	1.05	0.97	0.98
Reading span				
n	175	173	165	168
M (divided by 10)	1.88	1.89	1.89	1.82
SD (divided by 10)	0.91	0.97	0.77	0.83
Reading span (with outlier deleted; all raw scores < 80)				
n	175	172	165	168
M (divided by 10)	1.88	1.85	1.89	1.82
SD (divided by 10)	0.91	0.85	0.77	0.83

Note. MZ = monozygotic; DZ = dizygotic. With full sample, Reading Span, A [additive genetic influences] = .00; D [dominant/epistatic genetic influences] = .49; E [nonshared environmental influences] = .51. With one outlier (MZ twin with Reading Span = 80 deleted), A = .00; D = .54; E = .46.

correlations relative to MZ correlations, and dominance/epistasis will decrease DZ correlations relative to MZ correlations. Thus, we ran both ACE and ADE models, and used the model with the lowest AIC as our preferred model. For both digits forward and WRAT-3 reading, the ACE model was the preferred model. For both of these measures, shared environmental influences were statistically significant (based on the 95% confidence intervals [CIs]), and accounted for slightly less than one third of the variation. Additive genetic influence was statistically significant for WRAT-3 reading, accounting for nearly one half of the variation. Additive genetic influence on digits forward accounted for 19% of the variance overall, but was not statistically significant.

In contrast, for reading span, there was no evidence for shared environmental influence. The ADE model was the preferred model, and shared environmental influences were estimated at zero in the ACE model. In the ADE model, genetic influences were due entirely to dominance/epistasis factors, and these effects accounted for approximately one half the overall variance (51%, 95% CI = 0–60%). The 95% CIs for both additive (A) and dominance/epistasis (D) included zero, indicating that we did not have enough power to differentiate between A and D influences. The insufficient power was confirmed by the fact that a model without dominance/epistasis effects (the AE model, not shown) fit the data only marginally more poorly than did the full ADE model (LRC[1] = 2.92, p = .087). In the AE model, additive genetic influences were also estimated at approximately one half the overall variance (49%; 95% CI = 38%–58%). In contrast, a model without any genetic influence (the CE model, not shown) fit significantly more poorly than did the ACE model (LRC = 15.30, $p < .001$) and had the highest AIC value of all of the models (AIC = +14.92). Thus far, these results provide clear evidence for significant genetic influence on reading span, but whether the genetic effects are non-additive is less clear. However, comparison of the AIC values (not shown in Table 3) does suggest that the ADE model for reading span (AIC = −1.30) is a better model than the AE model (AIC = +1.62). Therefore, our multivariate analyses assume specific ACE influences on WRAT-3 Reading and Digit Span Forward, but specific ADE for reading span. Note that although we cannot decisively differentiate additive from nonadditive influences on reading span, retaining an ADE model does enable us to test for the presence of either or both types of genetic influence. For all measures, nonshared environmental influences (which can include measurement error) were significant, accounting for approximately 25% of the variation in WRAT-3 Reading and 50% of the variation in digits forward and reading span.

Multivariate genetic analyses (Table 4). We tested each of the three primary models (see Figure 1) under two different conditions, that is, based on the assumptions of ACE and ADE covariance, respectively. For all six models, we assumed specific ACE effects on Digits Forward and WRAT-3 Reading, and specific ADE influences on reading span (based on the univariate results). Although not shown here, we did formally test these assumptions in additional analyses (i.e., switching the specific C effect to a specific D effect for Digits Forward and WRAT-3 Reading, and switching the specific D influence to a

Table 2
Twin Correlations for Monozygotic and Dizygotic Twins From Saturated Model

Variable	Twin A Digit Span Forward	Twin A Reading span	Twin A WRAT-3	Twin B Digit Span Forward	Twin B Reading span	Twin B WRAT-3
Twin A Digit Span Forward		**.26**	**.40**	.39	*.18*	*.28*
Twin A Reading span	**.29**		**.38**	*.13*	.09	*.15*
Twin A WRAT-3 Reading	**.40**	**.34**		*.17*	*.16*	.51
Twin B Digit Span Forward	.50	*.31*	*.34*		**.27**	**.31**
Twin B Reading span	*.25*	.55	*.39*	**.29**		**.38**
Twin B WRAT-3 Reading	*.40*	*.37*	.78	**.35**	**.46**	

Note. WRAT-3 = Wide Range Achievement Test–3. Monozygotic twins are shown below the diagonal; dizygotic twins are shown above the diagonal. Phenotypic (within-person) correlations are presented in boldface type; cross-twin, within-trait correlations are underlined; cross-twin, cross-trait correlations are in italic.

Table 3
Proportion of Variance Explained and 95% Confidence Intervals (CIs) From Univariate Analyses

Model	−2 Log likelihood	*df*	AIC	Proportion of variance A	C	D	E
			Digit Span Forward				
ACE	**2,929.68**	**686**	**−7.64**	**.19**	**.31**	—	**.50**
95% CI				.00–.52	.02–.52	—	.40–.61
ADE	2,934.19	686	−3.12	.53	—	.00	.47
95% CI				.43–.61	—	.00–.24	.39–.57
			WRAT-3 Reading				
ACE	**1,756.63**	**682**	**−5.17**	**.49**	**.27**	—	**.24**
95% CI				.28–.74	.03–.46	—	.19–.30
ADE	1,761.28	682	−0.51	.76	—	.00	.24
95% CI				.55–.81	—	.00–.22	.19–.29
			Reading span				
ACE	1,683.83	677	1.62	.49	.00	—	.51
95% CI				.38–.58	.00–.14	—	.42–.62
ADE	**1,680.91**	**677**	**−1.30**	**.00**	—	**.51**	**.49**
95% CI				.00–.53	—	.00–.60	.40–.60

Note. AIC = Akaike's information criterion. A = additive genetic influences; C = shared environmental influences; D = dominance/epistatic genetic influences; E = nonshared environmental influences. WRAT-3 = Wide Range Achievement Test–3. Dashes indicate parameter set to zero. Degrees of freedom vary across measures because of missing data. The best-fitting model for each test is in boldface.

specific C influence for reading span). Switching the C or D effects resulted in models that fit significantly more poorly than did our original models, thereby confirming that these assumptions were accurate.[2]

Results from the primary multivariate genetic factor analyses revealed a number of patterns. First, on the basis of *p* values from the comparison of each primary model with the saturated model, both the independent pathways and the common pathways models fit the data well regardless of whether we start with ACE or ADE covariance (*p*s from .52 to .67). Second, under both conditions, the measurement model fit quite poorly (both *p*s $< .001$), and the LRC values from the comparison with the less restrictive common pathways model were also highly significant ($p < .001$). Third, there is no evidence that the common pathways model fit the data significantly more poorly than the less restrictive independent pathways model (*p*s = .66 and .85). In addition, it can be observed from Table 4 that the fit of the common pathways model was identical under the ACE and ADE covariance conditions. Inspection of the parameter estimates revealed that the 95% CIs for the additive genetic influence on variation in the underlying latent factor (A_C) included 1.0, indicating that covariance among our three measures could be accounted for entirely by additive genetic effects (see Figure 2). Not surprisingly, the A-only covariance model was the most parsimonious model (−2LL = 6,182.41, *df* = 2042, $p = .69$, AIC = −43.90). As such, Figure 2 presents the parameter estimates from the common pathways model with 100% of the covariance accounted for by additive genetic factors.

As can be seen in Figure 2, covariation among Digits Forward, reading span, and WRAT-3 Reading is accounted for by a single underlying latent factor, and all of the variation contributed by this factor is due to additive genetic influences. Figure 2 also shows that there are significant genetic and environmental influences on variation in each measure that are not accounted for by the latent phenotype. Consistent with most behavioral genetic studies, between about one quarter and one half of the variation in the three measures was attributed to nonshared environmental influences, and these nonshared environmental influences were measure-specific for all three tests. For WRAT-3 Reading, the specific genetic influence was not statistically significant (LRC = .01, $p = .94$), but dropping the specific shared environmental influence resulted in a significant deterioration in fit (LRC = 13.42, $p < .001$). Likewise, for digits forward, a model without specific shared environmental influences fit marginally more poorly than the full model (LRC = 2.83, $p = .09$), but a model without specific genetic influence fit nearly as well as the full model (LRC = 0.03, $p = .86$). Moreover, a model dropping both specific A and specific C influences on digits forward simultaneously resulted in a significant deterioration in fit, LRC(2) = 22.92, $p < .001$. Thus, for both Digits Forward and for WRAT-3 Reading, the source of specific familial influence appears to be shared environmental factors that account for approximately 25% of the overall variation in each measure. In contrast, the influence of genes specific to each measure is quite minimal, accounting for only 1–2% of the overall variation.

A very different pattern emerged for reading span. Our analyses suggested that the best-fitting model allowed for specific additive genetic and dominance/epistasis effects on variation in reading span and did not allow for specific shared environmental factors. Figure 2 shows that dominant/epistatic genetic effects specific to reading span accounted for an additional 23%

[2] Results of the additional analyses are available on request.

Table 4
Model Fitting Results for ACDE Covariance Among Cognitive Measures

Model	−2 Log likelihood	*df*	*p*	AIC	LRC[a]	*df*[a]	*p*[a]
Saturated	6,148.31	2003	—	—	—	—	—
ACE covariance							
Independent pathways	6,178.63	2036	.60	−35.69	—	—	—
Common pathways	6,181.02	2040	.67	−41.30	2.39	4	.66
Measurement model	6,269.45	2046	.001	35.14	88.43	6	<.001
ADE covariance							
Independent pathways	6,180.20	2036	.52	−34.11	—	—	—
Common pathways	6,181.54	2040	.65	−40.77	1.34	4	.85
Measurement model	6,274.01	2046	.001	36.70	92.47	6	<.001

[a] Because these are nested models, the common pathways model is compared with the independent pathways model, and the measurement model is then compared with the common pathways model. Dashes indicate that parameters are not applicable because the model is used as a comparison model. A = additive genetic influences; C = shared environmental influences; D = dominant/epistatic genetic influences; E = nonshared environmental influences; AIC = Akaike's information criterion; LRC = likelihood-ratio chi-square. All models allow for specific ACE influences on Digit Span Forward and Wide Range Achievement Test–3 Reading, and for specific ADE influences on Reading span. The LRC is based on comparisons with ACE and ADE models.

of variation, whereas additive genetic effects were estimated at zero. Although neither the additive genetic nor the dominance/epistasis parameters was statistically significant when tested individually (*p*s >.10), dropping both the additive (A) and non-additive (D) genetic influences simultaneously resulted in a significant deterioration in fit, LRC(2) = 13.42, $p < .001$, indicating that although there are clearly significant genetic influences on reading span that are not shared with genetic influences on the latent factor, we have inadequate power to determine conclusively whether these influences are additive or non-additive.

Heritability estimates from this model were .27 for digits forward, .51 for reading span, and .52 for WRAT-3 Reading (see Figure 2). For digits forward and WRAT-3 Reading, virtually all of the genetic influence on variation came from the genetic influences on the latent factor. Despite similar total heritabilities for

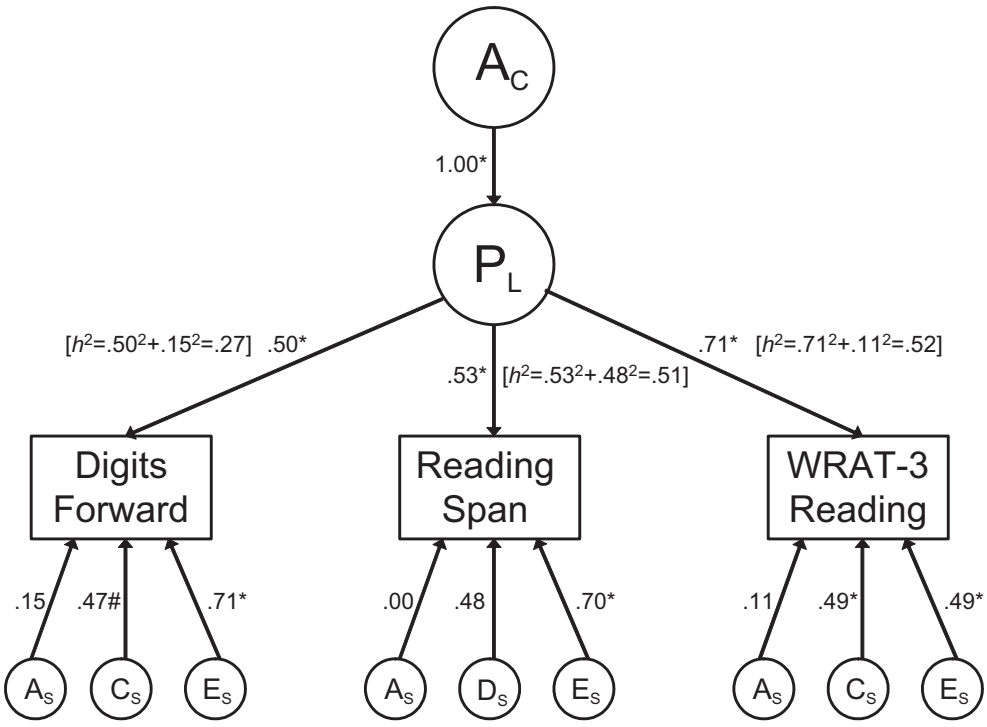

Figure 2. Standardized parameter estimates from common pathways model with 100% additive genetic covariance. To simplify the display, only one twin is represented in the diagram. A_C = additive genetic influences that are *common* to all three tests. A_S = additive genetic influences; C_S = shared environmental influences; D_S = dominant/epistatic influences; E_S = nonshared environmental influences that are *specific* to a particular test. P_L = underlying latent phenotype. h^2 = overall heritability. WRAT-3 = Wide Range Achievement Test–3. #$p < .09$. *$p < .05$.

reading span and WRAT-3 Reading, approximately one half of the genetic influence on variation in reading span came from genes that were specific to reading span.

Discussion

We sought to determine the extent to which common or specific genetic and environmental factors account for the component processes in reading span by modeling reading alone, short-term memory alone, and concurrent reading plus remembering. The best-fitting model was a common pathways A-only covariance model; that is, the covariation among the measures operates through a single latent phenotype that is influenced by additive genetic factors only. The model accounts for 100% of the variance in each measure, but the common factor accounted for about 50% of the variance in WRAT-3 Reading, about 25% of the variance in Digits Forward, and about 25% of the variance in reading span. These results are consistent with the modest observed phenotypic correlations among the three measures (*r*s = .26–.46), indicating that a substantial amount of variation in each of these measures was due to genetic or environmental factors that are not in common. For WRAT-3 Reading and Digit Span Forward, test-specific influences were due to both shared and unique environmental influences. Shared environmental influences may reflect factors such as parental education, parental socioeconomic status, or quality of schooling. This also means that these factors are not related to the cognitive ability that is tapped by all three measures simultaneously via the latent phenotype. On the other hand, about one half of the genetic influences on reading span were test-specific. None of the other tests had any significant test-specific genetic influences.

To interpret these results, we first consider what the latent phenotype represents. There is no short-term memory (storage/maintenance) component in WRAT-3 reading and no reading component in digits forward. We suggest that what is common to all three measures is likely to be general cognitive ability (*g*), or perhaps general verbal ability given that these were all verbal tests. The highest phenotypic correlation among the measures was $r = .46$, but the model shows that essentially all of their shared variance is accounted for by common genetic influences.

Most of the literature on complex span tasks suggests that the concurrent processing dimension in such tasks reflects an executive function component of verbal working memory. For example, a non-twin factor analysis by Engle and Oransky (1999) concluded that short-term/working memory tests consist of two factors: one representing the central executive and one reflective of storage alone. Reading span loaded on their central executive factor. Because reading span was the only measure requiring concurrent processing in our study, we further suggest that the test-specific genetic influences on reading span represent genes that are important for the executive component of working memory. However, we also acknowledge that other explanations are possible (e.g., Cowan et al., 2005). Cowan et al. (2005) noted that different working memory tasks were somewhat differentially correlated with particular aptitude tests. This suggests that different working memory tasks tap differing skill components. The tests in our study did not cover the entire set of skill components that may be involved in reading span performance. Therefore, the remaining executive component of this (or any) working memory task cannot necessarily be taken as a "pure" index of executive function, and the extent to which it may or may not be modality-specific remains to be determined.

In a multivariate twin analysis of complex working memory span tasks, Ando et al. (2001) also found a higher order general cognitive factor and evidence for specific genetic influences on the executive component of working memory. In our data, overall heritability for reading span was .51, with 55% of genetic variance coming from the common factor and 45% being specific to reading span. Ando et al. had a verbal working memory executive factor with corresponding values of .43, 47%, and 53%. For Digit Span Forward, these values were .27, 93%, and 7% in our study. Corresponding values for Ando et al.'s verbal working memory storage factor were .22, 95%, and 5%. On the surface, it appears that our results are quite consistent with the results of their verbal span test. However, these similarities could be somewhat coincidental because the two studies are not directly comparable for at least two reasons. First, Ando et al. were not able to fully differentiate putative storage and executive components because each of those indices was derived from performance during a concurrent processing task; in other words, they did not utilize any single ability tests (e.g., storage only). Their executive measure was sentence verification, and their storage measure was words recalled. What we inferred to reflect an executive component was based on that component of words recalled in reading span that was not in common with the other test measures, whereas Ando et al. considered words recalled to reflect storage ability. Second, they included general cognitive ability measures in their model and we did not; instead, a common (general ability) factor emerged in our model as a composite of the individual measures. Their inclusion of general ability measures is clearly of interest, but it is also indicative of the different aims of the two studies.

Several studies suggest that inhibition and interference control, which may be thought of as a family of inhibition-related executive functions, may be key abilities underlying reading span performance (Friedman & Miyake, 2004; May & Hasher, 1998; May et al., 1999; Whitney, Arnett, Driver, & Budd, 2001). After examining different types of inhibition/interference functions, Friedman and Miyake (2004) found that reading span performance was related to resistance to proactive interference, but not to inhibition of prepotent responses or resistance to external distractor interference.

Neuroimaging studies suggest that the neural substrate for resistance to proactive interference is strongly linked to prefrontal cortex. Activations associated with inhibition or interference have been observed in dorsolateral prefrontal cortex (Bunge, Ochsner, Desmond, Glover, & Gabrieli, 2001; Menon, Adleman, White, Glover, & Reiss, 2001), ventrolateral prefrontal cortex (D'Esposito, Postle, Jonides, & Smith, 1999; Menon et al., 2001; Pliszka et al., 2006; Schulz et al., 2005), and medial frontal/anterior cingulate (Botvinick, Cohen, & Carter, 2004; Kerns et al., 2004; Menon et al., 2001; Schulz et al., 2005). Activations in other brain regions have also been associated with inhibition or interference monitoring in these and other studies, but those regions are typically regions that are widely considered to be part of prefrontal/working memory circuitry and they are rarely observed in the absence of frontal lobe activations. Therefore, the most parsimonious explanation seems to be that this function reflects some component of prefrontal function.

Several potential predictors of reading span performance were not addressed in the twin models. Given the restricted age range in our sample, it is not surprising that age was uncorrelated with reading span performance ($r = .03$, *ns*). Processing speed (based on CPT reaction time) was also uncorrelated with reading span performance ($r = -.03$, *ns*); moreover, component processes of the Conners CPT include key processes that are not subsumed by reading span. Education, occupation, and AFQT scores were each modestly correlated with reading span performance (rs = .25, .23, and .26, respectively, ps $< .0001$); in a multiple regression analysis, they accounted for 11% of the variance in reading span performance, $F(3, 638) = 25.79$, $p < .0001$. These measures were not included in the genetic analyses because (a) the primary goal of the study was to examine the genetic architecture of cognitive subprocesses in reading span, not to determine predictors of level of performance, and (b) they are not subprocesses of reading span. It is true that our genetic analyses of reading span were limited to only two other tests, but the component processes of those tests were subsumed by reading span. We do not purport to have covered all of the cognitive subprocesses involved in reading span, but our approach does allow for more precise conclusions about the particular component processes that we have examined.

The Conners CPT is also a classic test for evaluating attention deficit hyperactivity disorder (ADHD). Previous work with children has shown that the covariance between reading disability (based on word recognition scores) and ADHD inattention symptoms is almost entirely accounted for by common genetic influences (Willcutt, Pennington, & DeFries, 2000). Common genetic influences between reading and attentional difficulties would suggest that any attentional difficulties would be reflected in our common latent phenotype rather than the genetic influences that are specific to reading span. Moreover, in the present adult sample, CPT perceptual sensitivity (d') was only modestly correlated with WRAT-3 Reading ($r = .12$) or reading span ($r = .15$; ps $< .003$). Also, the number of individuals with reading disability was likely reduced because of our exclusion of participants with WRAT-3 Reading scores below 70.

Although there are several strengths of this study, there are some limitations as well. Given the sample demographics, we do not know how generalizable the findings may be to women, racial/ethnic minorities, or younger or older age cohorts. We can be confident that there are genetic influences that are specific to reading span, but it must be acknowledged that we do not have sufficient power to be absolutely certain whether these are additive or non-additive effects. The univariate estimates suggested that all of the genetic influence on reading span was non-additive. However, comparison of the specific DE and specific AE models indicated that the total genetic influence was about the same and that the overall fits were not very different from one another. (Results for these models are available on request). On the other hand, as noted in the Results section, switching from D to C influences for reading span (i.e., switching from a dominance/epistasis to an additive genetic model) resulted in a model that fit significantly more poorly and in which the C effects were estimated at zero. Given these findings and the rationale that we have articulated, we therefore lean toward these specific genetic influences on reading span being non-additive. Finally, we were unable to address all possible component processes in this study. Future studies could directly assess specific executive functions purported to be important in reading span performance (particularly resistance to proactive interference) and the extent to which processing speed may play a role.

In summary, we found significant test-specific genetic effects on reading span in addition to common genetic influences on reading span, Digits Forward, and WRAT-3 Reading. The best-fitting model indicated that the common genetic influences were additive and that the test-specific genetic influences on reading span were most likely non-additive. We argued that the most parsimonious explanation of these specific genetic influences is that they reflect an executive–inhibitory component of reading span that is likely to be mediated by prefrontal cortex. Alterations in the structure or functioning of components of prefrontal neural circuitry have been implicated in normal aging and in several neuropsychiatric disorders. Therefore, our results may have useful implications for finding susceptibility genes for cognitive decline and cerebral changes associated with aging as well as neuropsychiatric conditions that are associated with dysfunctional prefrontal circuitry.

References

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

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience, 9,* 357–381.

Ando, J., Ono, Y., & Wright, M. J. (2001). Genetic structure of spatial and verbal working memory. *Behavior Genetics, 31,* 615–624.

Baddeley, A. D. (1986). *Working memory.* Oxford, England: Oxford University Press.

Baddeley, A. D., Bressi, S., Sala, D., Logie, R., & Spinnler, H. (1991). The decline of working memory in Alzheimer's disease: A longitudinal study. *Brain, 114,* 2521–2542.

Baddeley, A. D., & Della Sala, S. (1996). Working memory and executive control. *Philosophical Transactions of the Royal Society of London, 351B*(1346), 1397–1403.

Baddeley, A. D., & Logie, R. H. (1999). The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). New York: Cambridge University Press.

Bayliss, D. M., Jarrold, C., Gunn, D. M., & Baddeley, A. D. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology: General, 132,* 71–92.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences, 8,* 539–546.

Brooks, A., Fulker, D. W., & DeFries, J. C. (1990). Reading performance and general cognitive ability: A multivariate genetic analysis of twin data. *Personality and Individual Differences, 11,* 141–146.

Bunge, S. A., Klingberg, T., Jacobsen, R. B., & Gabrieli, J. D. (2000). A resource model of the neural basis of executive working memory. *Proceedings of the National Academy of Sciences USA, 97,* 3573–3578.

Bunge, S. A., Ochsner, K. N., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (2001). Prefrontal regions involved in keeping information in and out of mind. *Brain, 124,* 2074–2086.

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

Condray, R., Steinhauer, S. R., van Kammen, D. P., & Kasparek, A. (1996). Working memory capacity predicts language comprehension in schizophrenic patients. *Schizophrenia Research, 20,* 1–13.

Condray, R., Steinhauer, S. R., van Kammen, D. P., & Kasparek, A. (2002). The language system in schizophrenia: Effects of capacity and linguistic structure. *Schizophrenia Bulletin, 28,* 475–490.

Conners, K. (1992). *Conners' Continuous Performance Test: Manual.* New York: Multi-Health Systems.

Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51,* 42–100.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

D'Esposito, M., Postle, B. R., Jonides, J., & Smith, E. E. (1999). The neural substrate and temporal dynamics of interference effects in working memory as revealed by event-related functional MRI. *Proceedings of the National Academy of Sciences USA, 96,* 7514–7519.

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

Eisen, S. A., True, W. R., Goldberg, J., Henderson, W., & Robinette, C. D. (1987). The Vietnam Era Twin (VET) Registry: Method of construction. *Acta Geneticae Medicae et Gemellologiae, 36,* 61–66.

Engle, R. W., & Oransky, N. (1999). Multi-store versus dynamic models of temporary storage in memory. In R. J. Sternberg (Ed.), *The nature of cognition* (pp. 515–555). Cambridge, MA: MIT Press.

Finkel, D., Pedersen, N. L., McGue, M., & McClearn, G. E. (1995). Heritability of cognitive abilities in adult twins: Comparison of Minnesota and Swedish data. *Behavior Genetics, 25,* 421–431.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General, 133,* 101–135.

Fuster, J. M. (1989). *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe* (2nd ed.). New York: Raven Press.

Gayán, J., & Olson, R. K. (2003). Genetic and environmental influences on individual differences in printed word recognition. *Journal of Experimental Child Psychology, 84,* 97–123.

Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum & V. Mountcastle (Eds.), *Handbook of physiology—The nervous system: Vol. V. Higher functions of the brain* (pp. 373–417). Bethesda, MD: American Physiological Society.

Grafman, J., Jonas, B. S., Martin, A., Salazar, A. M., Weingartner, H., Ludlow, C., Smutok, M. A., & Vance, S. C. (1988). Intellectual function following penetrating head injury in Vietnam veterans. *Brain, 111,* 169–184.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). New York: Academic Press.

Hayakawa, K., Shimizu, T., Ohba, Y., & Tomioka, S. (1992). Risk factors for cognitive aging in adult twins. *Acta Geneticae Medicae et Gemellologiae, 41,* 187–195.

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

Hervey, A. S., Epstein, J. N., & Curry, J. F. (2004). Neuropsychology of adults with attention-deficit/hyperactivity disorder: A meta-analytic review. *Neuropsychology, 18,* 485–503.

Hollingshead, A. B. (1975). *Four-factor index of social status.* Unpublished manuscript, Yale University.

Johansson, B., Whitfield, K., Pedersen, N. L., Hofer, S. M., Ahern, F., & McClearn, G. E. (1999). Origins of individual differences in episodic memory in the oldest-old: A population-based study of identical and same-sex fraternal twins aged 80 and older. *Journals of Gerontology: Psychological Sciences and Social Sciences, 54B,* P173–P179.

Kendler, K. S., Heath, A. C., Martin, N. G., & Eaves, L. J. (1987). Symptoms of anxiety and depression in a volunteer twin population: The etiologic role of genetic and environmental factors. *Archives of General Psychiatry, 43,* 213–221.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., 3rd, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303,* 1023–1026.

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

Kremen, W. S., Jacobson, K. C., Xian, H., Eisen, S. A., Waterman, B., Toomey, R., Neale, M. C., Tsuang, M. T., & Lyons, M. J. (2005). Heritability of word recognition in middle-aged men varies as a function of parental education. *Behavior Genetics, 35,* 417–433.

May, C. P., & Hasher, L. (1998). Synchrony effects in inhibitory control over thought and action. *Journal of Experimental Pyschology: Human Perception and Perfromance, 24,* 363–379.

May, C. P., Hasher, L., & Kane, M. J. (1999). The role of interference in memory span. *Memory and Cognition, 27,* 759–767.

McArdle, J. J., & Goldsmith, H. H. (1990). Alternative common factor models for multivariate biometric analyses. *Behavior Genetics, 20,* 569–608.

Menon, V., Adleman, N. E., White, C. D., Glover, G. H., & Reiss, A. L. (2001). Error-related brain activation during a Go/NoGo response inhibition task. *Human Brain Mapping, 12,* 131–143.

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

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

Nichols, R. C., & Bilbro, W. C. J. (1966). The diagnosis of twin zygosity. *Acta Genetica et Statistica Medica, 16,* 265–275.

Pedersen, N. L., Plomin, R., Nesselroade, J. R., & McClearn, G. E. (1992). A quantitative genetic analysis of cognitive abilities during the second half of the life span. *Psychological Science, 3,* 346–353.

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

Pliszka, S. R., Glahn, D. C., Semrud-Clikeman, M., Franklin, C., Perez, R., III, Xiong, J., & Liotti, M. (2006). Neuroimaging of inhibitory control areas in children with attention deficit hyperactivity disorder who were treatment naive or in long-term treatment. *American Journal of Psychiatry, 163,* 1052–1060.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2001). *Behavioral genetics* (4th ed.). New York: Worth.

Plomin, R., Pedersen, N. L., Lichtenstein, P., & McClearn, G. E. (1994). Variability and stability in cognitive abilities are largely genetic later in life. *Behavior Genetics, 24,* 207–215.

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 1–90). Hillsdale, NJ: Erlbaum.

Salthouse, T. A. (1990). Working memory as a processing resource in cognitive aging. *Developmental Review, 10,* 101–124.

Schulz, K. P., Tang, C. Y., Fan, J., Marks, D. J., Newcorn, J. H., Cheung, A. M., & Halperin, J. M. (2005). Differential prefrontal cortex activation during inhibitory control in adolescents with and without childhood attention-deficit/hyperactivity disorder. *Neuropsychology, 19,* 390–402.

Seidman, L. J., Oscar-Berman, M., Kalinowski, A. G., Ajilor, O., Kremen, W. S., Faraone, S. V., & Tsuang, M. T. (1995). Experimental and clinical neuropsychological measures of prefrontal dysfunction in schizophrenia. *Neuropsychology, 9,* 481–490.

Stevenson, J., Graham, P., Fredman, G., & McLoughlin, V. (1987). A twin study of genetic influences on reading and spelling ability and disability. *Journal of Child Psychology and Psychiatry, 28,* 229–247.

Thompson, P. M., Cannon, T. D., Narr, K. L., van Erp, T., Poutanen, V. P., Huttunen, M., Lonnqvist, J., Standertskjold-Nordenstam, C. G., Kaprio, J., Khaledy, M., Dail, R., Zoumalan, C. I., & Toga, A. W. (2001). Genetic influences on brain structure. *Nature Neuroscience, 4,* 1253–1258.

Tsuang, M. T., Bar, J. L., Harley, R. M., & Lyons, M. J. (2001). The Harvard Twin Study of Substance Abuse: What we have learned. *Harvard Review of Psychiatry, 9,* 267–279.

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

Wadsworth, S. J., Fulker, D. W., & DeFries, J. C. (1999). Stability of genetic and environmental influences on reading performance at 7 and 12 years of age in the Colorado Adoption Project. *International Journal of Behavioral Development, 23,* 319–332.

Wechsler, D. (1987). *Manual for the Wechsler Memory Scale—Revised.* San Antonio, TX: Psychological Corporation.

Whitney, P., Arnett, P. A., Driver, A., & Budd, D. (2001). Measuring central executive functioning: What's in a reading span? *Brain and Cognition, 45,* 1–14.

Wilkinson, G. S. (1993). *WRAT-3: Wide Range Achievement Test Manual.* Wilmington, DE: Wide Range.

Willcutt, E. G., Pennington, B. F., & DeFries, J. C. (2000). Twin study of the etiology of comorbidity between reading disability and attention-deficit/hyperactivity disorder. *American Journal of Medical Genetics (Neuropsychiatric Genetics), 96,* 293–301.

Williams, L. J., & Holahan, P. J. (1994). Parsimony-based fit indices for multiple-indicator models: Do they work? *Structural Equation Modeling, 1,* 161–189.

Wingfield, A., Stine, E. A. L., Lahar, C. J., & Aberdeen, J. S. (1988). Does the capacity of working memory change with age? *Experimental Aging Research, 14,* 103–107.

Received July 15, 2005
Revision received November 27, 2006
Accepted January 8, 2007 ■