Journal of Gerontology: PSYCHOLOGICAL SCIENCES
2008, Vol. 63B, No. 2, P84–P91

Storage and Executive Components of Working Memory: Integrating Cognitive Psychology and Behavior Genetics in the Study of Aging

William S. Kremen,[1,2] Hong Xian,[3,4] Kristen C. Jacobson,[5] Lindon J. Eaves,[6] Carol E. Franz,[1] Matthew S. Panizzon,[1] Seth A. Eisen,[3,7] Andrew Crider,[8] and Michael J. Lyons[9]

[1]Department of Psychiatry and [2]Center for Behavioral Genomics, University of California, San Diego, La Jolla.
Departments of [3]Veterans Affairs and [4]Internal Medicine, Washington University, St. Louis, Missouri.
[5]Department of Psychiatry, University of Chicago, Illinois.
[6]Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond.
[7]Department of Psychiatry, Washington University, St. Louis, Missouri.
[8]Department of Psychology, Williams College, Williamstown, Massachusetts.
[9]Department of Psychology, Boston University, Massachusetts.

We combined experimental cognitive and behavior genetic methods to investigate storage and executive components of working memory in 663 middle-aged male twins. A single latent factor model indicated that digits forward (storage) and two-digit transformation (executive + storage) scores were influenced by the same genes. Additional executive demands in digit transformation appeared to increase the variance of individual genetic differences from 25% for digits forward to 48% and 53% for the digit transformation scores. Although it was not the best model, a two-factor model also provided a good fit to the data. This model suggested the possibility of a second set of genes specifically influencing the executive component. We discuss the findings in the context of research suggesting that new genetic influences come into play if demand continues to increase beyond a certain threshold, a threshold that may change with task difficulty and with age.

Key Words: Cognitive psychology—Twin–behavior genetic methods—Working memory.

OUR goal in the present study was to integrate experimental cognitive psychology and twin–behavior genetic methods as a strategy for gaining further understanding of cognitive aging. To date, virtually all twin studies of cognition have utilized what is basically a traditional neuropsychological approach (Bouchard & McGue, 2003; Finkel, Pedersen, McGue, & McClearn, 1995; Finkel, Pedersen, Plomin, & McClearn, 1998; Lessov-Schlaggar, Swan, Reed, Wolf, & Carmelli, 2006; Luciano et al., 2001; Wright et al., 2001). We refer to this as a "broad brushstroke" approach because the tests are broad and multidetermined; each individual test requires many different abilities, any one of which could account for performance deficits. Broad brushstroke or composite test indices may have high heritability but they are also likely to be influenced by many genes, only some of which may be important for cognitive aging.

Experimental cognitive psychology utilizes a more fine-grained approach, generally by attempting to decompose a test, that is, parse cognition, into component processes. The primary test of interest may itself be multidetermined, but the other tests that are analyzed are intended to represent cognitive subprocesses of the primary test of interest so that component processes of the primary test can be elucidated. This sort of unidirectionality is a key feature of this approach; that is, one test is subsumed by or is a necessary component of the other, but not the reverse. This approach may be contrasted with something like a test of the relationship between a working memory and an IQ test; because either may include components of the other, it cannot be concluded that one is a component of the other and not the reverse.

A goal of parsing cognition is to isolate cognitive processes. Integrating the approach of parsing cognition with the twin method might provide a way to isolate not only the component cognitive processes of a test that are more likely to be associated with cognitive aging, but also the genetic influences that are specific to the most age-relevant cognitive processes. Thus, it could potentially increase the genetic signal-to-noise ratio, which could, in turn, improve our ability to successfully utilize genotyping to identify particular genes associated with task performance.

We illustrate this approach by mapping experimental cognitive psychology methods for parsing cognition onto twin–behavioral genetic methods in order to examine storage and executive components of verbal working memory, and the genetic and environmental influences on these processes. In our application of this strategy, we utilized two relatively simple tasks, one of which consisted of a basic component process of the other. We are unaware of other twin studies that have included attempts to decompose cognitive tasks in this way.

Verbal working memory is often conceptualized according to a model in which the central executive regulates the modality-specific storage system known as the phonological loop (Baddeley, 1986, 1996). Nongenetically informative cognitive psychology studies have strongly suggested that the prefrontal cortex is essential for working memory and for functioning of the central executive (Baddeley & Della Sala, 1996; Bunge, Klingberg, Jacobsen, & Gabrieli, 2000; D'Esposito, Postle, Ballard, & Lease, 1999), and that executive and working memory deficits are often associated with aging (Hasher &

Zacks, 1988; Hedden & Park, 2001; Salthouse, Babcock, & Shaw, 1991; Wingfield, Stine, Lahar, & Aberdeen, 1988).

Multivariate twin analysis has revealed a latent executive control factor with an overall heritability of 79% in older adults; this factor included tests that tap a variety of very different cognitive abilities, including set shifting, psychomotor speed, interference control, strategic search, and working memory (Swan & Carmelli, 2002). Our approach in the present study was to more narrowly focus on genetic and environmental influences on component processes of verbal working memory specifically.

We examined the genetic and environmental influences on storage (maintenance) and executive (manipulation) components of verbal working memory in a middle-aged twin sample. Middle age still tends to be an understudied stage of life in aging research (Finch, 1991), but it constitutes an important prelude to the examination of changes that take place in later life. Our measure of working memory storage ability was forward digit span. The heritability of digit span has been estimated in several behavior genetic studies. Heritability of a trait is the proportion of phenotypic variance attributable to genetic variance. In middle-aged and older adults, the heritabilities for digit span (composite of digits forward and backward) have generally been moderate (0.34–0.66; see Finkel et al., 1995; Hayakawa, Shimizu, Ohba, & Tomioka, 1992; Plomin, Pedersen, Lichtenstein, & McClearn, 1994). Separate heritability estimates have been made for digits forward and backward in some studies of older adults; the estimates of 0.00 and 0.27 for digits forward were lower than estimates of 0.49 and 0.44 for digits backward (Johansson et al., 1999; Pedersen, Plomin, Nesselroade, & McClearn, 1992). Note that these were all univariate analyses, which means that they provide no information about the genetic architecture of the different tests or subtests.

The digit transformation task was the working memory measure that we used that included both a storage and an executive component (see the Methods section). Ando, Ono, and Wright (2001) estimated the heritability of an executive component within a concurrent processing verbal working memory task to be 0.43. Their approach was a very useful one, but it differs from ours in a key way: they attempted to examine storage and executive components in a concurrent processing working memory task. Our strategy was to include a separate storage-only measure because we think that, at some level, one cannot assess a person's simple storage ability while the person is performing a concurrent processing task. A recent review noted the relative lack of data on genetic and environmental influences on executive control in aging as a significant knowledge gap (Deater-Deckard & Mayr, 2005).

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). The VET Registry was created as a general resource for behavior genetic research. We drew the twins in the present study from 3,322 twin pairs (6,644 individuals) who were previously interviewed by telephone in the now-completed Harvard Drug Study (Tsuang, Bar, Harley, & Lyons, 2001). The present study was a study of vulnerability to alcoholism, but we did not screen or select participants from the Harvard sample on the basis of substance use or any other characteristics. We randomly selected them from the Harvard sample with one exception; those with service in Vietnam were not recruited. We excluded this subgroup to avoid recruitment conflicts with another VET Registry study being conducted at the same time.

We determined zygosity on the basis of questionnaire and blood-group methods (Eisen, Neuman, Goldberg, Rice, & True, 1989), with approximately 95% accuracy compared to DNA analysis (Nichols & Bilbro, 1966; Peeters, Van Gestel, Vlietinck, Derom, & Derom, 1998). One can find a complete description of the Registry's construction in previous publications (Eisen, True, Goldberg, Henderson, & Robinette, 1987; Henderson et al., 1990).

In the present study, we assessed 693 individuals; lengthy face-to-face assessments necessitated a much smaller sample than used in the Harvard study. If both members of a pair agreed to participate, we had the twins flown in from around the country for a daylong series of assessments at the University of California, Davis in Sacramento and at Harvard Medical School in Boston. We gave the participants their choice of study site. Participants were living throughout the United States, making this a national sample. Moreover, comparisons with U.S. census data indicate that the present sample and the larger VET Registry sample are similar to American men in their age range in terms of demographic and health characteristics (cf. Kremen et al., 2006).

The study was approved by the Institutional Review Boards at participating sites, and all twins gave written informed consent to participate. There were 176 monozygotic (MZ) and 169 dizygotic (DZ) pairs; we tested 181 pairs in Boston, 163 pairs in Sacramento, and 1 pair of twins 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 were unable to participate. The mean age of the sample 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; 97% graduated high school or obtained a GED, and 33% were college graduates; 98% of the participants were employed full time and 1.7% were employed part time.

Tests and Procedures

We administered tests in counterbalanced order across twin pairs, but always in the same order within twin pairs. This report focuses on two short-term memory tasks that involve the auditory presentation of numbers. Forward digit span served as an index of the storage component of working memory, because participants must simply maintain the numbers presented in working memory. The measure we used was the total score on the digit span forward portion of the revised Wechsler Adult Intelligence Scale–Revised (Wechsler, 1981); possible raw scores ranged from 0 to 14. Digit transformation includes both executive and storage components of working memory. On each trial, participants heard four digits presented over head-

phones at the rate of approximately 1 per second. After a pause of about 1 second, they then heard one of the following instructions: "add 3," "add 4," or "add 0." The instruction "write" followed another 1-second pause, and participants had to write the four numbers that resulted from adding that amount to each digit. They were not permitted to write the initially presented digits or to perform any written calculations. We presented the "add" instructions in pseudorandom order, with no more than two consecutive trials having the same instruction.

We presented 20 trials in total; we treated the first 2 as practice trials and the remaining 18 as the test trials. The add-0 trials require information storage only (similar to digits forward); virtually no errors were made in these trials. The add-3 and add-4 trials require both storage ability (retaining the digits presented) and executive ability (performing the mental calculations). The total number of correct responses on the 4 add-3 trials (possible raw score range: 0–16) and 5 add-4 trials (possible raw score range: 0–20) provided two indices of executive working memory. All participants were able to do the practice add-3 trial, indicating that they had sufficient mathematical ability to perform the task.

Statistical Analysis

The total add-3 and add-4 scores for digit transformation were negatively skewed. Therefore, we treated them as ordinal variables and used a threshold-based model that calculates the proportion of the sample that falls into each level. There were five levels for add-3 scores and six levels for add-4 scores. We performed statistical analyses with Mx (Neale, Boker, Xie, & Maes, 2003), a widely used maximum-likelihood-based program for the analysis of twin data. Although digits forward was approximately normally distributed, we recoded it as an ordinal variable with seven levels because it is not possible to model continuous and categorical variables simultaneously with the current Mx program. Fortunately, simulation studies have shown that when the ordinal variable has multiple categories, the analysis of ordinal data in twin models provides very similar estimates to the analysis of continuous data without significant loss of power (Neale, Eaves, & Kendler, 1994). We constrained the variance of ordinal data to equal unity in all analyses.

We could have combined add-3 and add-4 information into a single composite measure and performed a bivariate analysis, but that would not allow us to determine whether the nonshared genetic influences were entirely specific to one component or the other or split between the two. In contrast, our trivariate model does enable us to further decompose the covariances to address this question.

Having created ordinal variables, we report all correlations as polychoric correlations (although for comparison with other studies, we also provide the "raw" means and standard deviations). In order to examine the genetic and environmental influences on storage and executive components of working memory, we fit a trivariate twin model to the digits forward and digit transformation scores simultaneously. This trivariate model is an extension of the basic biometric so-called ACE model that is commonly used in twin research (Eaves, Last, Young, & Martin, 1978; Neale & Cardon, 1992), in which variation can be partitioned into variance that is due to additive genetic influences (A), shared or common environmental influences (C), and nonshared or individual-specific environmental influences (E). Shared environmental influences refer to environmental factors that contribute to twin similarity. Nonshared environmental influences refer to environmental factors that contribute to differences between twins. Measurement error is assumed to be random, that is, uncorrelated across twins; it is, therefore, also included in the nonshared environmental variance.

To assist in comparisons across models, we first fit a multivariate saturated model that captures the observed data perfectly and a Cholesky decomposition model that estimates the observed genetic and environmental variances and covariances but is agnostic about their origins. We then tested theoretically based models beginning with a two-factor model with one (storage) factor loading on all three measures and a second (executive) factor loading on add-3 trials and add-4 trials only. Next we tested whether a one-factor model could account for the data without a significant reduction in fit. Finally, we tested a reduced version of the best-fitting model to determine whether dropping nonsignificant parameters would result in a still more parsimonious model.

The one-factor common pathways model (Figure 1) assumes that an underlying latent phenotype is responsible for the covariation among the measures; in other words, genes and environment influence the correlation among variables by means of a common pathway (McArdle & Goldsmith, 1990). The paths (λs) from the latent phenotype to the individual measures 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. Under this model, there are also specific genetic, shared, and nonshared environmental influences on the three variables. The two-factor model has assumptions similar to the one-factor model, but it allows for two independent underlying latent phenotypes. Because our theoretically based two-factor model proposes a second latent phenotype that loads solely on the two-digit transformation variables, we constrained factor loadings (i.e., λ paths) for the second factor to be equal, in order to identify the two-factor model.

We fit all models to the raw data by means of maximum likelihood. We compared the fit of each submodel (indexed by the -2 log likelihood [-2LL]) to the relevant comparison model by means of the likelihood-ratio test (LRT) statistic and the Akaike Information Criterion (AIC; Akaike, 1987; Williams & Holahan, 1994). The LRT statistic is the difference in -2LL, distributed as a chi-square with degrees of freedom equal to the difference in degrees of freedom, between a comparison model and a nested (reduced) model. The reduced (more simple) model is generally accepted as the better model if the LRT is nonsignificant. The AIC provides an index of goodness of fit and parsimony; the more negative the AIC, the better the balance between the two. Therefore, if the LRT is nonsignificant for two or more competing models, then the model with the lowest AIC is considered to be the preferred model.

Results

There were 663 twins (95.7% of the sample) with valid digits forward and digit transformation scores (339 MZs; 324 DZs). The mean digits forward score was 8.18 ($SD = 2.14$;

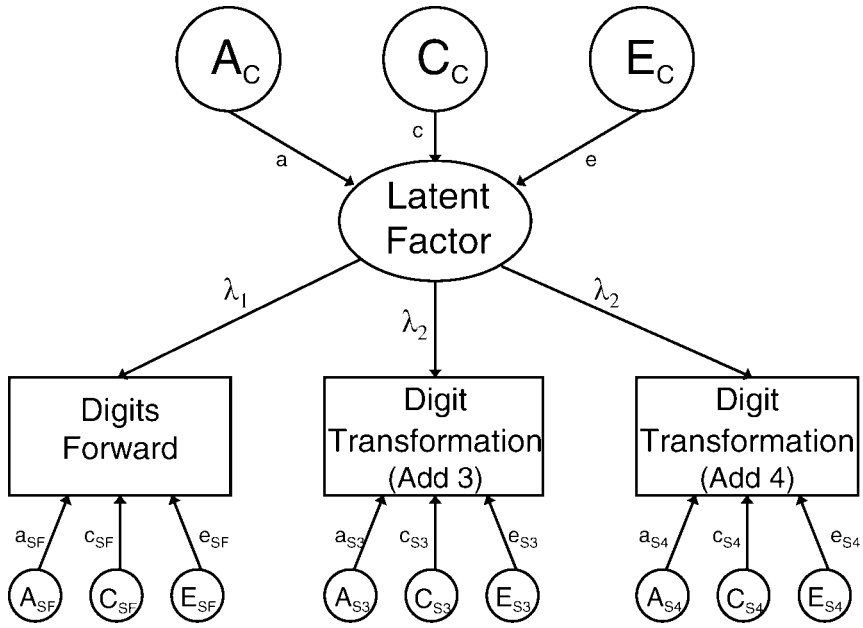

Figure 1. Final common pathways model: To simplify the display, only one twin is represented. A_C = additive genetic influences, C_C = shared environmental influences, and E_C = nonshared environmental influences that are *common* to all three tests. A_S = additive genetic influences, C_S = shared environmental influences, and E_S = nonshared environmental influences that are *specific* to a particular test. Lambda (λ) paths represent factor loadings for the three measured variables on the latent phenotype. Subscripts $_{F, 3,}$ and $_4$ indicate digits forward, digit transformation add 3, and digit transformation add 4, respectively.

range = 3–14) and the mean digit transformation scores were 14.28 (*SD* = 2.05; range = 4–16) for add-3 trials and 17.89 (*SD* = 2.49; range = 4–20) for add-4 trials. The MZ correlation (with 95% confidence interval, or CI) for digits forward was 0.59 (CI = 0.48–0.70) and the DZ correlation was 0.41 (CI = 0.27–0.55). The MZ correlation for digit transformation add-3 trials was 0.41 (CI = 0.26–0.56) and the DZ correlation was 0.21 (CI = 0.04–0.38). The MZ correlation for add-4 trials was 0.54 (CI = 0.41–0.67) and the DZ correlation was 0.16 (CI = 0.02–0.30).

The phenotypic correlation was 0.37 (CI=0.30–0.44) between digits forward and add-3 trials and 0.35 (CI=0.28–0.42) between digits forward and add-4 trials. The phenotypic correlation between add-3 and add-4 trials was 0.54 (CI = 0.51–0.54). As one would expect, these measures are positively correlated.

Table 1 shows the fit of the different models. The Cholesky (Model 2) provided a good fit to the data compared with the fully saturated model (Model 1). The two-factor model (Model 3) did not fit the data more poorly than did the less restrictive Cholesky model, and it had a more negative AIC value than Model 2. The one-factor model (Model 4) did not result in a significantly worse fit to the data than the two-factor model, and it had a more negative AIC value than did the two-factor model. In the reduced one-factor model (Model 5), all of the nonsignificant parameters from Model 4 were set to zero. All but one of those parameters (the E influences on the latent factor) were already estimated at zero in the full model. This reduced model had the best overall balance between goodness of fit and parsimony.

Standardized parameter estimates from the full one-factor common pathways model are shown in Figure 2. This model is essentially an AE model because additive genetic influences accounted for 92% of the variance in the latent factor and individual-specific environmental influences accounted for the remaining 7% of variance (total does not equal 100% due to rounding error). There were no significant genetic influences

Table 1. Model-Fit Statistics

	Absolute Fit Statistics			Relative Fit Statistics		
Model	−2LL	*df*	AIC	LRT	Δ*df*	*p*
1. Fully Saturated	6359.06	1912	0			
2. Cholesky	6433.24	1969	−39.82	74.18	57	.06
3. Two-Factor Model	6435.44	1970	−39.62	2.20	1	>.10
4. One-Factor Model	6441.02	1973	−40.04	5.58	3	>.10
5. One-Factor Model (reduced)	**6442.72**	**1980**	**−52.34**	**1.70**	**7**	**>.10**

Note: AIC = Akaike Information Criterion; −2LL = −2 log likelihood; LRT = likelihood ratio test statistic; Δ*df* = difference in degrees of freedom between the two models being compared. The LRT is based on nested comparisons of models: Model 2 is compared with Model 1; Model 3 is compared with Model 2; Model 4 is compared with Model 3; Model 5 is compared with Model 4. Best model is indicated in boldface.

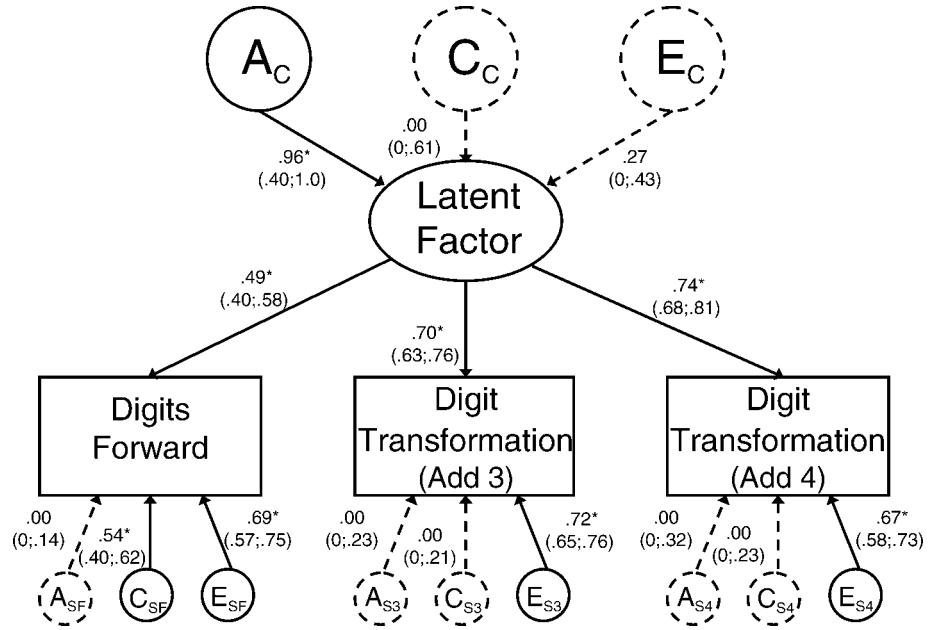

Figure 2. Standardized parameter estimates from the full one-factor common pathways model (Model 4 in Table 1; note that figure elements are explained in Figure 1). One calculates the standardized variance components (i.e., the proportion of variation in each variable that is due to genetic, shared, and nonshared environmental factors) by summing the square of the product of the λ coefficients and the standardized common A, C, or E parameters plus the square of the standardized specific A, C, or E parameters. Heritabilities are calculated as follows: $h^2 = (0.49 \times 0.96)^2 + (0.00)^2 = 0.22$ for digits forward; $h^2 = (0.70 \times 0.96)^2 + (0.00)^2 = 0.45$ for add-3 trials; $h^2 = (0.74 \times 0.96)^2 + (0.00)^2 = 0.50$ for add-4 trials (for numerals with an asterisk, $p < .05$). Dashed lines indicate parameters that were dropped in the reduced model (Model 5 in Table 1).

beyond those accounted for by the latent factor. Specific shared environmental factors accounted for 29% of the variance in digits forward. Individual-specific environmental influences accounted for a larger proportion of variance (45%–53%) in all three measures as compared with shared environmental factors. In this model, heritabilities were 0.22 (CI = 0.15–0.65) for digits forward, 0.45 (CI = 0.13–0.54) for add-3 trials, and 0.50 (CI = 0.41–0.58) for add- 4 trials (see Figure 2).

Even though the reduced one-factor model was the most parsimonious, we did not depict it in a separate figure because it would be largely redundant with Figure 2. The reduced model is an A-only covariance model in which the covariance among the three measures is accounted for solely by genetic influences (C and E influences on the latent factor were dropped). There was little change in the parameter estimates. Additive genetic influences accounted for 92% of the variance in the full one-factor model and 100% of the variance in the reduced model. Specific shared environmental influences on digits forward and individual-specific environmental influences on each measure accounted for similar amounts of variance as in the full model. Heritabilities in the reduced model were 0.25 (CI = 0.17–0.33) for digits forward, 0.48 (CI = 0.38–0.55) for add-3 trials, and 0.53 (CI = 0.44–0.62) for add-4 trials.

Although our analyses indicate that the one-factor common pathways model was a better model by AIC than the two-factor model, we did choose to illustrate the two-factor model (see Figure 3) because it may have useful implications regarding the genetic architecture of storage and executive components of working memory. This model suggested that additive genetic influences accounted for 34% (0.58^2) of the variance in the second (executive) factor, but this value was not statistically significant based on the 95% CI.

Discussion

We examined the genetic and environmental architecture underlying the relationship between storage and executive components of working memory. The most parsimonious model (Model 5) was a reduced one-factor common pathways model. This model indicated that 100% of the genetic variation in the measures is shared and that environmental influences did not account for any of the covariance among the measures. In the full one-factor model, nonshared environmental influences accounted for only 7% of variance and additive genetic influences accounted for 92% of the variance in the common factor. In either case, all or virtually all of the variance was accounted for by additive genetic influences, and there were no genetic influences that were specific to any of the three individual measures.

One explanation of this finding would be that both the storage and executive components are influenced by the same set of genes. A second explanation could be that the common genetic factor reflects only working memory storage capacity, and variation in the executive component of working memory is due to nonshared environment only. These explanations are supported by the fact that storage capacity is the only component common to all three measures. Only digit transformation requires executive ability, and nonshared environmental factors were the only influence outside of the common genetic factor that contributed to variation in the digit transformation measures.

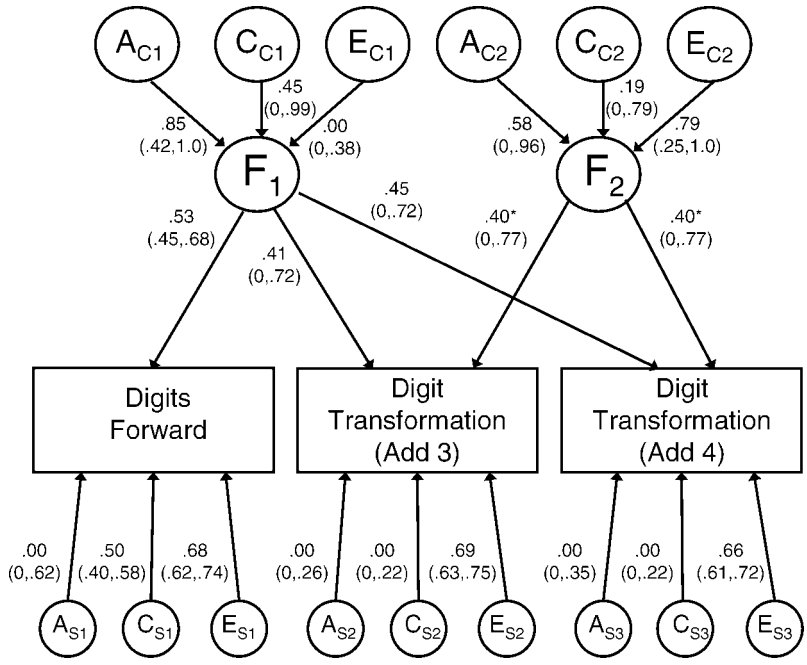

Figure 3. Standardized parameter estimates from the two-factor model (Model 3 in Table 1; note that figure elements are explained in Figure 1 and standardized variance components are explained in Figure 2). F_1 = storage factor; F_2 = executive factor.

Findings from other studies suggest that these explanations are unlikely. Ando and colleagues (2001) did find evidence for genetic influences that were specific to the executive components of their working memory tasks, although their approach was different in that they examined concurrent processing only; that is, they utilized tasks that required concurrent storage and executive ability but did not separately measure performance on tasks that required storage only. In the same sample as the present study, we also found evidence for specific genetic influences on an executive component of working memory based on tasks that required reading alone, short-term memory alone, and concurrent reading and short-term memory (Kremen et al., in press). Therefore, the idea that executive ability in working memory is influenced only by individual-specific environmental factors, or that the same genes account for both the storage and executive components of working memory, seems unlikely.

These possibilities also seem unlikely given the results of the two-factor model in the present study. Those results, which did not yield the most parsimonious model but still provided a good fit to the data, were at least suggestive of a separate set of genetic influences that is specific to the executive component of working memory (see Figure 3). The findings from our other study (Kremen et al., 2007) and that of Ando and associates (2001) raise the question of why the two-factor model in the present study did not indicate significant genetic influences on the executive factor.

It appears that we may have been underpowered to detect genetic influences on executive ability that are not correlated with storage capacity. With a larger sample, the genetic influences on the executive factor in our two-factor model would likely be significant. Nevertheless, in Kremen and colleagues (2007), we did find evidence for specific genetic influences on an executive component of working memory in these same participants with the same sample size. These different outcomes suggest that reduced power in the present analysis is more likely to be a function of differences in the measures used. The add-3 and add-4 scores were based on only four and five trials, respectively. A bivariate analysis based on digits forward and the total digit transformation score combining add-3 and add-4 scores did yield similar results, although we did not present that analysis for reasons noted in the Methods section. Even so, the combined measure was based on only nine trials, raising that possibility that digit transformation measures based on a larger number of trials may have had less error that could have resulted in the genetic influences on the executive factor being significant.

Perhaps the most likely explanation of our results is that we were underpowered as a result of the nature of the executive function task. In other words, the genetic architecture in the present data might be similar to that of other working memory tests such as those in the studies of Ando and colleagues (2001) or Kremen and associates (2007), *if* one takes a threshold effect into account. According to this explanation, genes influencing executive ability are the same genes that influence storage capacity, but only up to a point; new genetic effects are expressed when cognitive demand surpasses a specific threshold. Although we are unaware of behavioral genetic studies that have addressed this question with respect to cognition, results from a twin study of cardiovascular response (CVR) during exercise are consistent with this explanation. van den Bree, Schieken, Moskowitz, and Eaves (1996) found evidence for two sets of genetic factors: one factor that influenced variation in CVR at rest and continued with decreasing impact during exercise, and a separate factor that influenced variation in CVR at the start of exercise and continued to increase. Increasing

load over the range of submaximal exercise did not result in the expression of totally new genetic effects.

It may be that digit transformation remains below some cognitive load threshold, and thus does not result in or require genetic influences that are additional to those contributing to digits forward. Or, as suggested by our two-factor model, digit transformation may have only enough cognitive load for its specific genetic influences to manifest themselves rather weakly. Indeed, compared with digit transformation, the tasks used by Ando and colleagues (2001) and in our previous analyses (Kremen et al., 2007) to assess executive function in working memory were more cognitively demanding concurrent processing tasks. This is borne out by comparison of the distributions and proportion of individuals achieving maximum scores on the reading span test in our previous analysis (Kremen et al.) and on the digit transformation measures. The concurrent processing tasks may, therefore, exceed a certain threshold that necessitates the expression of additional, new genetic factors in order to achieve adequate performance. Such a threshold might be analogous to either the transition from rest to exercise or to going above "the range of submaximal exercise" in the CVR study. One could test this hypothesis directly by extending a similar research strategy to that of the present study in order to test the effects of systematically increasing the cognitive load of the tests being administered.

In any case, our results also indicate that the genetic influence on individual differences in working memory is stronger for executive ability than for storage capacity. In both the full and the best-fitting reduced models, the heritabilities of add-3 and add-4 were roughly double the heritability of digits forward. Post hoc analyses indicated that constraining all three heritabilities to be equal resulted in a significant deterioration in fit (results available upon request). This pattern is consistent with the notion that the additional cognitive demand of the digit transformation tests "spreads out" or magnifies the individual genetic differences that are observed on the less demanding digits forward measure; the more demanding cognitive tasks seem to exaggerate the same genetic effects. Interestingly, van den Bree and colleagues (1996) argued that their results supported this explanation for CVR; they observed a similar pattern in that the relative importance of genetic factors on variation in CVR during exercise increased at higher levels of exertion despite the fact that there was no evidence for new genetic effects at the higher levels.

Some functional neuroimaging studies may also suggest neuroanatomical parallels to this pattern. In some paradigms, the same brain regions that are activated during a more simple task also manifest increased activations when working memory or executive function demands are increased; a more demanding task does not necessarily call new brain regions into play (Bunge et al., 2000; Gold, Berman, Randolph, Goldberg, & Weinberger, 1996).

As we noted in the introduction, executive function and working memory are cognitive functions that generally involve prefrontal cortex and are susceptible to age-related decline. Several neuroimaging studies have indicated that older adults manifest increased activation or recruit additional brain regions compared with younger adults, or that cognitive tasks that do not require recruitment of frontal-executive function in young adults often do require it in older adults (DiGirolamo et al., 2001; Grady, Springer, Hongwanishkul, McIntosh, & Winocur, 2006; Reuter-Lorenz & Mikels, 2005). These patterns are consistent with theories suggesting that older adults must compensate for reduced cognitive capacity or reduced neural efficiency (Cabeza, 2002; Reuter-Lorenz & Mikels). Therefore, a task such as digit transformation might, in effect, become more demanding as these middle-aged individuals get older and require more effortful resource allocation. If so, a longitudinal twin study—such as we are planning in our current Vietnam Era Twin Study of Aging (VETSA; Kremen et al., 2006)—might be expected to reveal that the same task requires additional genes for individuals beyond a certain age to be able to maintain performance levels.

Although the demographic characteristics of our sample may limit our ability to generalize findings to women or ethnic minorities, our findings are informative about genetic and environmental influences on cognitive processes in midlife. Two very common approaches to the study of aging are to conduct cross-sectional comparisons of old and young adults, and longitudinal studies beginning at roughly age 65 years. One potential shortcoming of these strategies is that we do not learn anything about changes that take place between young adulthood and older age. Late middle age is likely to be a key transition phase that for many people will be just prior to or on the cusp of significant age-related cognitive changes. By focusing on this period, it may enable us to observe functioning both before and after important changes occur. This approach makes good practical as well as scientific sense because it offers the possibility of studying important changes without requiring an inordinately long follow-up interval. Whether we think in terms of a goal of identifying early indicators of cognitive decline or of learning how people can maintain better functioning for a greater portion of their life spans (longer "health spans"), it will be important to learn more about this key period for cognitive change.

Acknowledgments

Preparation of this article was supported in part by the National Institute of Alcohol Abuse and Alcoholism under Grant AA10586 and the National Institute on Aging under Grants AG18386, AG18384, AG22381, and AG22982. The U.S. Department of Veterans Affairs has provided financial support for the development and maintenance of the VET Registry. Numerous organizations have provided invaluable assistance in the conduct of this study: 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; and 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

Address correspondence to William S. Kremen, PhD, Department of Psychiatry, Center for Behavioral Genomics, University of California–San Diego, 9500 Gilman Drive (MC 0738), La Jolla, CA. E-mail: wkremen@ucsd.edu

References

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

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: Oxford University Press.

Baddeley, A. D. (1996). The fractionation of working memory. *Proceedings of the National Academy of Sciences USA, 93,* 13,468–13,472.

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

Bouchard, T. J., Jr., & McGue, M. (2003). Genetic and environmental influences on human psychological differences. *Journal of Neurobiology, 54,* 4–45.

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.

Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: The HAROLD model. *Psychology and Aging, 17,* 85–100.

D'Esposito, M., Postle, B. R., Ballard, D., & Lease, J. (1999). Maintenance versus manipulation of information held in working memory: An event-related fMRI study. *Brain and Cognition, 41,* 66–86.

Deater-Deckard, K., & Mayr, U. (2005). Cognitive change in aging: Identifying gene-environment correlation and nonshared environment mechanisms. *Journal of Gerontology: Psychological Sciences, 60B,* P24–P31.

DiGirolamo, G. J., Kramer, A. F., Barad, V., Cepeda, N. J., Weissman, D. H., Milham, M. P., et al. (2001). General and task-specific frontal lobe recruitment in older adults during executive processes: A fMRI investigation of task-switching. *Neuroreport, 12,* 2065–2071.

Eaves, L. J., Last, K. A., Young, P. A., & Martin, N. G. (1978). Model-fitting approaches to the analysis of human behavior. *Heredity, 41,* 249–320.

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.

Finch, C. E. (1991). Middle-age: An evolving frontier in gerontology. *Neurobiology of Aging, 12,* 1–2.

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.

Finkel, D., Pedersen, N. L., Plomin, R., & McClearn, G. E. (1998). Longitudinal and cross-sectional twin data on cognitive abilities in adulthood: The Swedish Adoption/Twin Study of Aging. *Developmental Psychology, 34,* 1400–1413.

Gold, J. M., Berman, K. F., Randolph, C., Goldberg, T. E., & Weinberger, D. R. (1996). PET validation of a novel prefrontal task: Delayed response alternation. *Neuropsychology, 10,* 3–10.

Grady, C. L., Springer, M. V., Hongwanishkul, D., McIntosh, A. R., & Winocur, G. (2006). Age-related changes in brain activity across the adult lifespan. *Journal of Cognitive Neuroscience, 18,* 227–241.

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.

Hedden, T., & Park, D. (2001). Aging and interference in verbal working memory. *Psychology and Aging, 16,* 666–681.

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.

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. *Journal of Gerontology: Psychological Sciences, 54B,* P173–P179.

Kremen, W. S., Jacobsen, K. C., Xian, H., Eisen, S. A., Eaves, L. J., Tsuang, M. T., et al. (2007). Genetics of verbal working memory processes: A twin study of middle-aged men. *Neuropsychology.*

Kremen, W. S., Thompson-Brenner, H., Leung, Y. J., Grant, M. D., Franz, C. E., Eisen, S. A., et al. (2006). Genes, environment, and time: The Vietnam Era Twin Study of Aging (VETSA). *Twin Research and Human Genetics, 9,* 1009–1022.

Lessov-Schlaggar, C. N., Swan, G. E., Reed, T., Wolf, P. A., & Carmelli, D. (2006). Longitudinal genetic analysis of executive function in elderly men. *Neurobiology of Aging, 28,* 1759–1768.

Luciano, M., Wright, M. J., Smith, G. A., Geffen, G. M., Geffen, L. B., & Martin, N. G. (2001). Genetic covariance among measures of information processing speed, working memory, and IQ. *Behavior Genetics, 31,* 581–592.

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

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2003). *Mx: Statistical modeling* (6th 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.

Neale, M. C., Eaves, L. J., & Kendler, K. S. (1994). The power of the classical twin study to resolve variation in threshold traits. *Behavior Genetics, 24,* 239–258.

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.

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.

Reuter-Lorenz, P. A., & Mikels, J. A. (2005). The aging mind and brain: Implications of enduring plasticity for behavioral and cultural change. In P. Baltes, P. A. Reuter-Lorenz, & F. Roesler (Eds.), *Lifespan development and the brain: The perspective of biocultural co-constructivism* (pp. 255–278). London: Cambridge University Press.

Salthouse, T. A., Babcock, R. L., & Shaw, R. J. (1991). Effects of adult age on structural and operational capacities in working memory. *Psychology and Aging, 6,* 118–127.

Swan, G. E., & Carmelli, D. (2002). Evidence for genetic mediation of executive control: A study of aging male twins. *Journal of Gerontology: Psychological Sciences, 57*(2), P133–P143.

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.

van den Bree, M. B., Schieken, R. M., Moskowitz, W. B., & Eaves, L. J. (1996). Genetic regulation of hemodynamic variables during dynamic exercise. The MCV twin study. *Circulation, 94,* 1864–1869.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale—Revised.* San Antonio, TX: Psychological Corporation.

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.

Wright, M., De Geus, E., Ando, J., Luciano, M., Posthuma, D., Ono, Y., et al. (2001). Genetics of cognition: Outline of a collaborative twin study. *Twin Research, 4,* 48–56.

Received November 28, 2006
Accepted September 10, 2007
Decision Editor: Thomas M. Hess, PhD