Published in final edited form as:

Neuropsychology. 2011 July ; 25(4): 488–498. doi:10.1037/a0022569.

Genetic Architecture of Learning and Delayed Recall: A Twin Study of Episodic Memory

Matthew S. Panizzon, Ph.D.¹, Michael J. Lyons, Ph.D.², Kristen C. Jacobson, Ph.D.³, Carol E. Franz, Ph.D.¹, Michael D. Grant, Ph.D.², Seth A. Eisen, M.D.⁴, Hong Xian⁴, and William S. Kremen, Ph.D.^{1,5,6}

¹Department of Psychiatry, University of California, San Diego

²Department of Psychology, Boston University

³Department of Psychiatry, University of Chicago

⁴Research Services, St. Louis VA Medical Center

⁵Center for Behavioral Genomics, University of California, San Diego

⁶San Diego VA Health Care System, San Diego

Abstract

Objective—Although episodic memory is often conceptualized as consisting of multiple component processes, there is a lack of understanding as to whether these processes are influenced by the same or different genetic determinants. The aim of the present study was to utilize multivariate twin analyses in order to elucidate the degree to which learning and delayed recall, two critical measures of episodic memory performance, have common or different genetic and environmental influences.

Method—Participants from the Vietnam Era Twin Study of Aging (314 monozygotic twin pairs, 259 dizygotic twin pairs, and 47 unpaired twins) were assessed using the second edition of the California Verbal Learning Test (CVLT-2). Mean age at the time of the evaluation was 55.4 years ($sd = 2.5$).

Results—Model fitting revealed the presence of a higher-order latent factor influencing learning, short- and long-delay free recall, with a heritability of .36. The best-fitting model also indicated specific genetic influences on learning, which accounted for 10% of the overall variance. Given that learning involves the acquisition and retrieval of information, whereas delayed recall involves only retrieval, we conclude that these specific effects are likely to reflect genes that are specific to acquisition processes.

Conclusion—These results demonstrate that even in non-clinical populations, it is possible to differentiate component processes in episodic memory. These different genetic influences may have implications for gene association studies, as well as other genetics studies of cognitive aging and disorders of episodic memory such as Alzheimer's disease or mild cognitive impairment.

Correspondence: Dr. Matthew S. Panizzon, Department of Psychiatry, University of California, San Diego, 9500 Gilman Drive (MC 0738), La Jolla, CA 92037-0738; Tel: 858-534-8269; Fax: 858-822-5856; mspanizz@ucsd.edu.

The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at www.apa.org/pubs/journals/neu

Keywords

Episodic Memory; Twin Study; Genetics; Aging

Episodic memory, the ability to accurately recall details from one's daily experiences, is an essential cognitive process that is heavily impacted by normal cognitive aging, and is associated with a host of neurological and psychiatric conditions (Airaksinen, Larsson, Lundberg, & Forsell, 2004; Butters, Delis, & Lucas, 1995; Cirillo & Seidman, 2003; Mariani, Monastero, & Mecocci, 2007; Nilsson, 2003; Petersen et al., 1999; Schaie, 1996; Zec, 1993). Like most measures of cognitive functioning, episodic memory appears to be heritable (i.e., a significant proportion of its variance is attributable to genetic factors), with the majority of twin studies conducted to date reporting estimates of its genetic influence in the range of .30 to .60 (Alarcon, Plomin, Fulker, Corley, & DeFries, 1998; T. J. Bouchard, Segal, & Lykken, 1990; Finkel, Pedersen, & McGue, 1995; Johansson et al., 1999; McClearn et al., 1997; Pedersen, Plomin, Nesselroade, & McClearn, 1992; Swan et al., 1999; Thapar, Petrill, & Thompson, 1994; Volk, McDermott, Roediger, & Todd, 2006). Yet despite the significant degree to which behavior genetics researchers have studied episodic memory, there remain a number of fundamental aspects of its underlying genetic structure that have yet to be explored. Among these is the genetic relationship between learning and delayed recall.

Experimental and clinical perspectives of episodic memory have long viewed the construct as a complex set of processes that include the ability of the individual to initially attend to, encode, and then later retrieve information (Baddeley, 2004; Butters et al., 1995; Erickson & Scott, 1977; Luria, 1973; Squire, Knowlton, & Musen, 1993; Tulving, 1983). On tests of episodic memory these processes can be represented by measures of learning ability, which capture both the acquisition and immediate retrieval of information, and measures of delayed recall, which capture only retrieval processes. The relationship between learning and delayed recall has proven to be critical in the assessment of memory disorders, for instance helping to distinguish between encoding and retrieval deficits (Butters et al., 1995). However, in non-clinical populations the two measures have been found to correlate strongly with one another and have frequently been shown to load onto the same higher-order factor (Delis, Freeland, Kramer, & Kaplan, 1988; Delis, Jacobson, Bondi, Hamilton, & Salmon, 2003; Millis, Malina, Bowers, & Ricker, 1999). Such findings have led to debate as to whether learning and delayed recall capture different cognitive constructs or whether they are actually redundant measures of the same process (Delis et al., 2003; Nee, Berman, Moore, & Jonides, 2008; Ranganath & Blumenfeld, 2005).

From a genetic perspective, strong phenotypic relationships provide little information about the degree of shared genetic influences between measures. For instance, a recent study by our group demonstrated that while measures from a test of planning and problem-solving, the Tower of London, loaded onto a single phenotypic factor, examination of the genetic influences revealed the presence of two genetic factors, one accounting for speed and the other for efficiency (Kremen et al., 2009). Thus, although learning and delayed recall are highly correlated in non-clinical populations, it is unclear whether this means that they share the same underlying genetic factors. Clarification of this relationship may provide valuable insights into the processes that are involved in episodic memory, as well as the variety of conditions that result in episodic memory impairment. Determining whether learning and delayed recall are, at least in part, influenced by non-shared genetic factors would provide support for the view that these are indeed non-redundant aspects of episodic memory, and establish the presence of distinct biological influences that are not ordinarily observable. This may in turn improve the ability of researchers to identify processes, either biological or

environmental, that drive age-related changes in memory, as well as changes observed in AD and MCI. If, however, these measures share all of their genetic influences then models of how episodic memory changes over the course of normal aging or disease progression may have to be reconsidered. With respect to gene association studies, identifying the presence of unique genetic influences for either learning or delayed recall would enable researchers to refine their phenotypes of interest by filtering out the effects of variables that contribute superfluous genetic influences, ultimately leading to a more targeted search for the genes involved.

To date, behavior genetics studies of episodic memory, unlike the experimental and clinical traditions, have been based largely on single measures of performance and have consequently ignored the relationship between learning and delayed recall. In order to fully elucidate the genetic relationship between these measures a multivariate application of the classical twin design is required. Whereas univariate twin analyses decompose the variance of a phenotype of interest into genetic and environmental components, multivariate twin analyses allow for the further decomposition of the covariance between multiple phenotypes (Neale & Cardon, 1992). As a result, the degree to which observed relationships are driven by genetic and/or environmental influences can be determined, as can the degree to which latent genetic and environmental influences of the phenotypes are predictive of one another (i.e. genetic and environmental correlations). Multivariate twin analyses also allow for genetically informative factor models to be tested, models that allow for sources of common genetic and environmental influences to be constrained into simplified structures while simultaneously capturing the degree to which the phenotypes of interest are influenced by unique, or variable-specific, genetic and environmental influences.

While no multivariate twin analysis of learning and delayed recall has been conducted to date, one prior twin study does warrant mention. Utilizing a sample of twins in their early to mid-seventies, Swan and colleagues (1999) subjected data from the first edition of the California Verbal Learning Test (CVLT) to a phenotypic principal components analysis, and subsequently determined the heritability of the resulting factor scores. Measures of learning and delayed recall loaded strongly onto the same factor, labeled as verbal learning and recall, the heritability of which was determined to be .56 (Swan et al., 1999). These results suggest a common source of genetic influences between learning and delayed recall; however, by examining the heritability of a predefined factor, the authors were not able to determine whether unique genetic influences were present for the underlying measures. Factors that are defined at the phenotypic level through the examination of standard covariance estimates typically account for only a portion of the overall variance in the variables of interest. The residual variance for these variables is far from irrelevant, and can be used in multivariate twin analyses to identify variable-specific genetic and environmental influences that are crucial for distinguishing between measures that are highly correlated at the phenotypic level. Multivariate twin methods make it possible to examine what is going on “beneath” the phenotypic level and are, thus, a powerful approach that may be able to shed new light on these episodic memory components.

In the present study, we utilized multivariate twin analyses in order to elucidate the genetic and environmental relationships between learning and delayed recall on the second edition of the California Verbal Learning Test (CVLT-2) (Delis, Kramer, Kaplan, & Ober, 2000). Learning was defined as the total number of words recalled across the five initial learning trials of the test, while delayed recall was examined in terms of performance on both short- and long-delay free recall conditions. We hypothesized that based on the role of retrieval processes in each of these measures, as well as previous evidence of robust phenotypic correlations, a significant degree of genetic overlap would exist between our measures of learning and delayed recall. Furthermore, we made three predictions regarding the possible

presence of unique genetic influences. If all components of episodic memory are influenced by a single, global source of genetic influences, then model testing would indicate the presence of no variable-specific genetic influences. If there are some genetic influences that are specific for acquisition and retrieval, then model testing would result in residual genetic influences for the learning trials versus short- and long- delay recall. If there are some specific genetic influences as a function delay interval, then model testing would indicate residual genetic influences on each of the three memory measures.

Methods

Participants

Data were collected as part of wave 1 of the Vietnam Era Twin Study of Aging (VETSA), a longitudinal study of cognitive aging with a baseline in midlife (Kremen et al., 2006). All participants in the VETSA were recruited from the Vietnam Era Twin (VET) Registry, a nationally distributed sample of male-male twin pairs who served in the United States military at some point between 1965 and 1975. Detailed descriptions of the VET Registry's composition and method of ascertainment have been reported elsewhere (Eisen, True, Goldberg, Henderson, & Robinette, 1987; Henderson et al., 1990). Zygosity for 92% of the VETSA sample was determined by analysis of 25 satellite markers obtained from blood samples. For the remaining 8% of the sample zygosity was determined through a combination of questionnaire and blood group methods. A comparison of these two approaches within the VETSA sample has demonstrated an agreement rate of 95%.

In total, 1237 individuals participated in wave 1 of the VETSA. Participants were given the option of traveling to either the University of California San Diego or Boston University for a day-long series of cognitive, psychological, and physiological assessments. In rare cases (<3% of participants) project staff traveled to the participant's home town to complete the evaluation. Informed consent was obtained from all participants at the start of the assessment session. The present analyses were based on data from 1194 participants: 323 monozygotic (MZ) twin pairs, 250 dizygotic (DZ) twin pairs, and 48 unpaired twins. Participant data were omitted from the present analyses if a valid administration of the CVLT-2 was not obtained, or if the participant reported a history of brain disease or brain damage.

Measures

The CVLT-2 was administered to all VETSA participants in the context of a larger neurocognitive battery. In accordance with the standard administration procedures, a list of 16 words was read to the participant over five trials. After each trial the participant was asked to repeat all of the words he could remember. Following these initial learning trials an interference list was read, and the participant was then asked to recall all of the words from the new list as possible. The interference list was promptly followed by a short delay free recall of the first list. Approximately 20 minutes later there was a long delay free recall of the first list. Directly following the short and long delay free recall conditions, a cued recall condition was administered in which the participant was prompted with the semantic categories of the words on the first list. The total of trials 1 through 5 was utilized as our measure of learning ability. The short and long delay free recall conditions were both utilized as indicators of delayed recall performance.

Statistical Analysis

In univariate twin analyses the total variance of any trait is decomposed into the proportion attributed to additive genetic (A) influences, shared or common environmental (C) influences (i.e., environmental factors that make members of a twin pair more similar to one

another), and unique environmental (E) influences (i.e., environmental factors that make twins different from one another, including measurement error). The resulting model is commonly referred to as the ACE model (Eaves, Last, Young, & Martin, 1978; Neale & Cardon, 1992). Additive genetic influences are assumed to correlate 1.0 between monozygotic (MZ) twins because they share 100% of their DNA. Dizygotic (DZ) twins share on average 50% of their segregating DNA, and are therefore assumed to correlate .50 for additive genetic influences. The shared environment is assumed to correlate 1.0 between both members of a twin pair, regardless of their zygosity (i.e., the equal environments assumption). It is also assumed that the means and variances for each variable are not significantly different between MZ and DZ twins. In the present study we are able to test the latter, but not the former of these assumptions; however, the fact that heritability estimates for cognition tend to be similar in twins reared apart or together suggests that violation of the equal environments assumption is not likely to be a significant threat to the validity of the present analyses (Bouchard, 1998; Bouchard & McGue, 2003). In addition, it should be noted that the present analyses do not take into account the potential presence of either gene-environment interaction or gene-environment correlation, and as a result the latent genetic and environmental influences are assumed to be independent of one another (Jinks & Fulker, 1970). Should such interactions or correlations indeed be present, the resulting estimates of genetic and environmental influences may vary in the context of different environmental factors.

In order to elucidate the genetic and environmental relationships between the CVLT variables, we fit a series of multivariate twin models. These analyses extend the univariate ACE model so as to further decompose the covariance between traits (i.e., a correlation that would be observed in a non-genetic analysis) into genetic and environmental components. We first fit a multivariate Cholesky decomposition to the data (see Figure 1a) in order to estimate the magnitude of genetic and environmental covariance between the variables of interest. These covariance estimates were then be used to calculate genetic and environmental correlations. The genetic correlation (r_G) between any two phenotypes (i and j) is defined as their genetic covariance (A_{ij}) divided by the square root of the product of their individual genetic variances (A_i and A_j).

$$r_G = \frac{A_{ij}}{\sqrt{(A_i \times A_j)}}$$

Note that this definition is the same as the correlation with which non-genetics researchers are familiar except that it is based only on the genetic variances and covariance. Environmental correlations, both shared and unique, are calculated in the same way utilizing the respective variance and covariance estimates. Conceptually, genetic and environmental correlations represent the degree to which genetic and environmental influences of one phenotype are predictive of the influences for another phenotype (Carey, 1988).

In addition to the multivariate Cholesky we fit independent pathway/biometric factors models (McArdle & Goldsmith, 1990) and common pathway/psychometric factors models (Kendler, Heath, Martin, & Eaves, 1987) in order to determine the degree to which genetic and environmental influences of immediate and delayed recall stem from the same higher-order factor, as well as whether variable-specific genetic and environmental influences were present. The independent pathway model (see Figure 1b) imposes higher-order genetic and environmental factors on the respective covariance estimates, while still allowing variable-specific genetic and environmental influences. These higher-order latent factors then act on each variable through separate, independent pathways. In the common pathway model (see Figure 1c), higher-order genetic and environmental influences act directly on a latent

phenotype, and then on the variables of interest through a common pathway that is equivalent to a factor loading. Like the independent pathway model, the common pathway model allows for the estimation of variable-specific genetic and environmental influences. When these specific influences are zero, all the genetic and environmental variance for a variable is assumed to come from the latent factor. Since unique environmental influences also include measurement error, the specific E influences on each variable cannot be set to zero. The independent pathway models were used to separately test for the presence of genetic and environmental factors without imposing constraints on the other sources of covariance. The common pathway model was then used to test for the presence of a global memory factor.

Evaluation of the goodness-of-fit for all models was performed using the likelihood-ratio chi-square test (LRT), which was calculated as the difference in the $-2 \log$ likelihood ($-2LL$) of the model in question relative to that of a fully saturated model. The saturated model recaptures the means, variances, and covariances of the data perfectly, and served as a standard against which each genetically informative model was compared. The LRT is distributed as a chi-square with degrees of freedom equal to the difference in the number of parameters between the two models. Non-significant LRT values ($p > .05$) indicate that a model does not result in a significant change in fit relative to the saturated model, and can therefore be considered as an accurate representation of, or good fit to, the data. We also used the Akaike Information Criterion (AIC) as an additional indicator of goodness-of-fit (Akaike, 1987). Calculated as the LRT value minus twice the difference in degrees of freedom, smaller AIC values represent a better balance on the part of the model between goodness-of-fit and parsimony.

All analyses were conducted using the raw data option of the maximum-likelihood based structural equation modeling software Mx (Neale, Boker, Xie, & Maes, 2004). Prior to model fitting, the CVLT variables were adjusted for the small but significant effect of age, correlations with which ranged from $-.09$ for long delay free recall to $-.12$ for the total of trials 1 through 5. The resulting residual scores were standardized to a mean of 0 and a variance of 1.0 in order to simplify the specification of start values and the appropriate ranges for model parameters.

Results

Descriptive statistics for the VETSA sample, as well as means and standard deviations for the CVLT-2 variables utilized are presented in Table 1. There were no differences between the MZ and DZ groups with respect to age, education, employment status, or ethnicity. Similarly, the performance of both groups on the CVLT measures of learning and delayed recall was equivalent. On average, the performance of VETSA participants on the CVLT-2 was consistent with that of the published normative sample (Delis et al., 2000).

Results from the trivariate Cholesky model are presented in Table 2. The Cholesky was a good fit relative to the fully saturated model ($p = .119$), indicating that several fundamental assumptions of the twin design were satisfied; for example, equality of means and variances within and across twins and zygosity groups (Neale & Cardon, 1992). Phenotypic correlations between the three measures were all significant, ranging from $r = .74$ for the relationship between the total of trials 1 through 5 and long delay recall, to $r = .83$ for the relationship between two delayed recall measures. All three phenotypes were found to be significantly heritable, with genetic influences accounting for 36% of the variance in the total of trials 1 through 5, 24% of the variance in short delay recall, and 35% of the variance in long delay recall. Shared environmental influences were noticeably lower than their genetic counterparts, ranging from 8% to 15% of the variance, with none reaching statistical

significance. Unique environmental influences accounted for the majority of the variance in all three measures. The genetic correlations between all three measures were highly significant, ranging from .87 to .99, and in each case the 95% confidence interval contained an upper limit of 1.0. Shared environment correlations were equally large, but did not reach statistical significance due to the non-significant contribution of C to the individual variances. Unique environmental correlations were moderate in comparison, ranging from .62 to .71, and were all significant.

All multivariate model fitting results are presented in Table 3. Following the trivariate Cholesky, we fit a series of models in which either a genetic or an environmental factor was imposed while the remaining elements were left unconstrained. This allowed us to separately test for the presence of genetic, shared environment, and unique environment factors. Each model resulted in a non-significant reduction in fit relative to the saturated model. The subsequent full independent pathway model in which genetic, shared environment, and unique environment factors were simultaneously imposed on the data also resulted in a non-significant reduction in model fit. The standardized parameter estimates for the full independent pathway model are presented in Figure 2. The higher-order latent genetic factor accounted for approximately 91% of the genetic variance across all three variables. Both short delay free recall and long delay free recall were found to have no variable-specific genetic influences (i.e., both parameters were equal to zero), while the total of trials 1–5 possessed some, albeit small and non-significant, residual genetic variance. The higher-order shared environment factor accounted for approximately 96% of the shared environment variance in the three variables, with small variable-specific influences present for the total of trials 1–5 and long delay free recall. The higher-order unique environment factor accounted for roughly 67% of the unique environment variance in the three factors, with significant residual unique environment influences present for each variable.

Like the full independent pathway model, the common pathway model was a good fit relative to the saturated model; however, the more negative AIC value for the common pathway model indicated a superior balance between model fit and parsimony. When tested directly against the full independent pathway model, the common pathway model resulted in a non-significant reduction in model fit ($LRT = 3.702$, $\Delta DF = 4$, $p = .448$), further suggesting that it provided a superior representation of the data. Figure 3 presents the standardized parameter estimates for the full common pathway model. The higher-order latent factor (P_L) was found to have a significant heritability of .36, and a non-significant shared environmental influence of .15. The latent phenotype accounted for 67% of the variance in the total of trials 1 through 5, 85% of the variance in short delay free recall, and 81% of the variance in long delay free recall. Of the residual variance components, only the unique environmental influences were found to be statistically significant based on 95% confidence intervals. Genetic and shared environment variance components accounted for roughly 1% of the variance or less with the noted exception of the residual genetic influences on the total of trials 1 through 5, which was found to account for 8.4% of the phenotypic variance.

Fixing at zero all parameters in Figure 3 that accounted for approximately 1% of the variance or less (i.e., all variable-specific shared environmental influences, as well as the variable-specific genetic influences for short and long delay free recall) further improved the fit of the common pathway model and resulted in the lowest AIC value of all the preceding models (-35.473). In this reduced common pathway model (see Figure 4) the variable-specific genetic influences for the total of trials 1 through 5 now accounted for 10% of the variance and were statistically significant based on the 95% confidence intervals. Variable-specific unique environmental influences accounted for 17% to 23% of the overall variance, and were once again statistically significant. There was no change in the magnitude of the

genetic, shared environment, or unique environmental influences of the latent phenotype. In order to determine if all the genetic and shared environmental influences in the CVLT measures could be accounted for by the higher-order latent phenotype we fit a final measurement model to the data. The measurement model was similar to the full common pathway model except that all variable-specific genetic and shared environmental influences were constrained to zero, essentially eliminating them from the model, and only unique environmental influences were allowed to act on each variable independently of the latent phenotype. As shown in Table 3, the measurement model proved to be a poor fit relative to the saturated model. Moreover, when compared directly against the full CP model, the measurement model resulted in a significant change in model fit ($LRT = 24.10$, $\Delta DF = 6$, $p < .001$). Overall, the reduced common pathway model proved to be the best-fitting model for the data.

In order to determine if these results represent merely the global effects of general verbal ability, we repeated our analyses after adjusting each CVLT-2 variable for performance on the vocabulary subtest of the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999). Controlling for general verbal ability had minimal impact on the previously described results. Heritability estimates remained relatively consistent: .30 for the total of trials 1 through 5, .19 for short delay free recall, and .28 for long delay free recall. Phenotypic, genetic, and environmental correlations between the variables also remained consistent, as did the genetic and environmental factor structures. As was the case with our original analyses, a reduced common pathway model, one in which all variable-specific shared environmental influences as well as variable-specific genetic influences for short and long delay free recall were fixed at zero, provided the best representation of the data. The resulting higher-order latent factor was found to have a heritability of .34 and a shared environment influence of .12. Variable-specific genetic influences for the total of trials 1 through 5 accounted for 9% of variance, and were once again statistically significant. Variable-specific unique environmental influences accounted for 21% to 28% of the variance in their respective phenotypes. The measurement model was again a poor fit relative to the saturated model.

Discussion

In the present study we examined the degree to which measures of learning and delayed recall, as assessed by the CVLT-2, were influenced by shared as well as variable-specific genetic and environmental influences. Our analyses revealed a highly significant degree of genetic and environmental overlap among these measures. In addition, the best-fitting multivariate model revealed a single higher-order latent factor that accounted for the majority of the phenotypic variance in all of the variables examined. This factor possessed a heritability of .36 (i.e., 36% of the variance was accounted for by latent genetic influences), as well as shared and unique environmental influences of .15 and .49, respectively. At the specific variable level, significant residual genetic influences were observed for only the total of trials 1 through 5, our measure of learning ability. These genetic influences, which by definition were independent of the common genetic factor, accounted for 10% of the variance in the learning measure. A measurement model in which the variable-specific genetic and shared environmental variance components were constrained to be zero, forcing all genetic and shared environmental influences to originate from the latent phenotype, resulted in a significant worsening of model fit. Given that learning involves the acquisition and retrieval of information, whereas delayed recall involves only retrieval, we conclude that the observed variable-specific genetic effects for the total of trials 1 through 5 are likely to reflect genes that are specific to acquisition processes.

As conceptualized by the authors of the CVLT-2 (Delis et al., 2000), the total of trials 1 through 5 represents an individual's global verbal learning ability and is a reflection of auditory attention as well as the degree to which information can be immediately stored and added to over repeated trials, all processes that are critical to the acquisition of information. That this variable was found to possess significant residual (i.e., variable-specific) genetic influences, whereas the short and long delay recall conditions did not, lends strong support to the argument that learning does involve some neurobiological processes that are distinct from retrieval. There was also clear evidence for variable-specific unique environment effects on each of the CVLT measures. In our best-fitting model residual unique environmental influences accounted for 17% to 23% of the variance in each of the measures, and in each case the values were significant based on 95% confidence intervals. These variable-specific unique environmental influences are likely to represent the effects of measurement error on each of the phenotypes. The excellent test-retest reliability of the CVLT-2 measures we used, which ranged from .81 to .88 (Delis et al., 2000), is certainly consistent with this conclusion. As for the other source of unique environmental influences (E_L) the list of potential environmental factors that may influence performance on the CVLT is extensive, and includes level of education, history of head injury, as well as unexpected effects of the individual administering the test (Wiens, Tindall, & Crossen, 1994).

To date, a limited number of studies have examined the effects of specific candidate genes on distinct aspects of episodic memory, specifically on measures of immediate and delayed memory performance. Measures of immediate recall typically represent an individual's ability to reproduce information immediately following its presentation, and like our measure of learning ability involve acquisition and retrieval processes. Genetic variation in the *5-HT2a* receptor, a critical receptor for the functioning of serotonin within the brain, has been found to be associated with delayed but not immediate recall (de Quervain et al., 2003). Three studies have examined the effects of the *KIBRA* gene on immediate and delayed recall; however, these results are somewhat inconsistent, providing evidence for an association with delayed recall but not immediate (Papassotiropoulos et al., 2006), immediate recall but not delayed (Schaper, Kolsch, Popp, Wagner, & Jessen, 2007), and both processes simultaneously (Almeida et al., 2008). The $\epsilon 4$ allele of the Apolipoprotein E (*APOE*) gene has also been shown to have discrepant effects on aspects of episodic memory. Using data from the first edition of the CVLT, Bondi and colleagues (1999) found no effect of the *APOE* $\epsilon 4$ allele on the total of trials 1 through 5; however, a significant association was found with long delay free recall, as well as a trend level effect for short delay free recall (Bondi, Salmon, Galasko, Thomas, & Thal, 1999).

The results from these gene association studies are largely consistent with our present findings suggesting that some genes influencing acquisition are distinct from those influencing retrieval. Indeed, the findings of the present study help to make sense of why one might observe different genetic associations for learning (or immediate recall) and delayed recall. However, our results also suggest considerable genetic overlap, so that the genetic association findings may not always suggest unique relationships. Inconsistent results from gene association studies may stem from the fact that genes rarely exert their influence in isolation from other genes, or environmental factors for that matter; therefore, discrepancies in the literature may be the result of failing to appropriately account for gene-gene or gene-environment interaction (Moore & Williams, 2002). It may also be the case that the degree to which learning and delayed recall are influenced by shared genetic and environmental influences changes as a function of increasing age. Longitudinal twin studies have shown that over time new latent genetic and environmental influences on memory performance emerge, and subsequently contribute to the observed changes in memory performance later in life (McArdle & Plassman, 2009; Reynolds, Finkel, Gatz, & Pedersen, 2002). Although these studies have all conceptualized episodic memory as a single variable,

their results could imply that later in life new genetic factors begin to exert specific influences on either learning or delayed recall.

We must acknowledge some potential limitations of the present study. First, the all-male relatively homogenous nature of the VETSA sample limits our ability to generalize these findings to female populations or ethnic minorities. Similarly, it is unclear whether the present results will replicate in child or adolescent populations due to the substantial degree of cognitive and brain development that is likely to be taking place. These data are of course cross-sectional and provide no insight into whether the observed relationships are long standing, or whether they are the result of age-related changes. Longitudinal data are currently being collected as part of wave 2 of the VETSA, and will allow for the examination of how the genetic and environmental influences of learning and recall change as these late midlife men enter later life. Because information can only be recalled if it has first been learned, it is possible that the degree of genetic overlap between our measures of learning and delayed recall was somewhat inflated. If that were the case, however, it would also strengthen the conclusion that there were genetic influences that were specific to learning and independent of recall. Lastly, there remain aspects of episodic memory that we are unable to examine with our current study design. Clinical and experimental memory testing typically involves delays of anywhere from 20 to 40 minutes, however, there is evidence from functional and structural neuroimaging studies to suggest that the retrieval of information over longer delay periods (e.g., days or weeks) may utilize distinct neurocognitive processes that are not accessed with traditional testing paradigms (Takashima et al., 2006; Walhovd et al., 2006). Consolidation and retrieval over an extended period of time could, therefore, represent yet another aspect of episodic memory that may or may not be genetically related to the constructs examined in the present study.

Delis and colleagues have argued that the use of phenotypic shared-variance procedures (i.e., data analytic techniques that are based on the correlations between variables) may obscure otherwise distinct cognitive constructs when used on data collected from normal populations (Delis et al., 2003). Although multivariate twin analyses are essentially an elaborate form of a shared-variance procedure, we believe that the unique properties of this method, specifically the ability to decompose the covariance as well as residual variance into genetic and environmental components, allows for insights into cognition that are beyond the scope of phenotypic shared-variance procedures. In previous studies of neurocognitive processes we have demonstrated the ability of multivariate twin analyses to isolate the genetic influences specific to executive functioning on tests of verbal working memory (Kremen et al., 2007; Kremen et al., 2008), as well as differentiate genetic influences for speed and efficiency on a test of planning and problem solving (Kremen et al., 2009). In some cases, specific effects or different factors were observed even when there was only a single phenotypic factor. As with the present results, these findings were produced using samples of cognitively normal participants and with measures that possessed a substantial amount of method variance, two features that have been argued to be potential confounds for differentiating cognitive constructs in non-patient populations (Delis et al., 2003). Thus, the examination of the genetic relationships between variables does appear capable of differentiating cognitive processes when phenotypic shared-variance techniques fail to do so. Future studies will ultimately need to determine whether the genetic relationships observed in the present study remain stable or change later in life. Furthermore, focusing on how the genetic and environmental relationships between learning and delayed recall change or stay the same may enhance our ability to find specific determinants of cognition, cognitive aging, and aging-related disease processes—such as mild cognitive impairment or Alzheimer’s disease—that substantially affect episodic memory.

Acknowledgments

The VETSA project is supported by National Institutes of Health/National Institute on Aging (NIH/NIA) Grants R01 AG018386-06A2, R01 AG018384-06A2, R01 AG22381, and R01 AG 22982. The U.S. Department of Veterans Affairs has provided support for the development and maintenance of the Vietnam Era Twin Registry. Numerous organizations have provided invaluable assistance, including VA Cooperative Studies Program; Department of Defense; National Personnel Records Center, National Archives and Records Administration; the Internal Revenue Service; National Institutes of Health; National Opinion Research Center; National Research Council, National Academy of Sciences; the Institute for Survey Research, Temple University; Schulman, Ronca, and Bucuvalas, Inc. Most importantly, we gratefully acknowledge the cooperation and participation of the members of the Vietnam Era Twin Registry and their families. Without their contribution this research would not have been possible.

References

- Airaksinen E, Larsson M, Lundberg I, Forsell Y. Cognitive functions in depressive disorders: evidence from a population-based study. *Psychological Medicine*. 2004; 34:83–91. [PubMed: 14971629]
- Akaike H. Factor analysis and AIC. *Psychometrika*. 1987; 52:317–332.
- Alarcon M, Plomin R, Fulker DW, Corley R, DeFries JC. Multivariate path analysis of specific cognitive abilities data at 12 years of age in the Colorado Adoption Project. *Behavior Genetics*. 1998; 28:255–264. [PubMed: 9803018]
- Almeida OP, Schwab SG, Lautenschlager NT, Morar B, Greenop KR, Flicker L, et al. KIBRA genetic polymorphism influences episodic memory in later life, but does not increase the risk of mild cognitive impairment. *Journal of Cellular and Molecular Medicine*. 2008; 12:1672–1676. [PubMed: 18194457]
- Baddeley, AD. The psychology of memory. In: Baddeley, AD.; Kopelman, MD.; Wilson, BA., editors. *The Essential Handbook of Memory Disorders for Clinicians*. Hoboken, NJ: John Wiley & Sons; 2004.
- Bondi MW, Salmon DP, Galasko D, Thomas RG, Thal LJ. Neuropsychological function and apolipoprotein E genotype in the preclinical detection of Alzheimer's disease. *Psychology and Aging*. 1999; 14:295–303. [PubMed: 10403716]
- Bouchard TJ. Genetic and environmental influences on adult intelligence and special mental abilities. *Human Biology*. 1998; 70:257. [PubMed: 9549239]
- Bouchard TJ Jr, McGue M. Genetic and environmental influences on human psychological differences. *Journal of Neurobiology*. 2003; 54:4–45. [PubMed: 12486697]
- Bouchard TJ, Segal NL, Lykken DT. Genetic and environmental influences on special mental abilities in a sample of twins reared apart. *Acta Geneticae Medicae et Gemellologiae*. 1990; 39:193–206.
- Butters N, Delis DC, Lucas JA. Clinical assessment of memory disorders in amnesia and dementia. *Annual Review of Psychology*. 1995; 46:493–523.
- Carey G. Inference about genetic correlations. *Behavior Genetics*. 1988; 18:329–338. [PubMed: 3219111]
- Cirillo MA, Seidman LJ. Verbal declarative memory dysfunction in schizophrenia: from clinical assessment to genetics and brain mechanisms. *Neuropsychological Review*. 2003; 13:43–77.
- de Quervain DJ, Henke K, Aerni A, Coluccia D, Wollmer MA, Hock C, et al. A functional genetic variation of the 5-HT_{2A} receptor affects human memory. *Nature Neuroscience*. 2003; 6:1141–1142.
- Delis DC, Freeland J, Kramer JH, Kaplan E. Integrating clinical assessment with cognitive neuroscience: construct validation of the California Verbal Learning Test. *Journal of Consulting and Clinical Psychology*. 1988; 56:123–130. [PubMed: 3346437]
- Delis DC, Jacobson M, Bondi MW, Hamilton JM, Salmon DP. The myth of testing construct validity using factor analysis or correlations with normal or mixed clinical populations: lessons from memory assessment. *Journal of the International Neuropsychological Society*. 2003; 9(6):936–946. [PubMed: 14632252]
- Delis, DC.; Kramer, JH.; Kaplan, E.; Ober, BA. *California Verbal Learning Test-Second Edition*. San Antonio, Texas: The Psychological Corporation; 2000.

- Eaves LJ, Last KA, Young PA, Martin NG. Model-fitting approaches to the analysis of human behavior. *Heredity*. 1978; 41:249–320. [PubMed: 370072]
- Eisen SA, True WR, Goldberg J, Henderson W, Robinette CD. The Vietnam Era Twin (VET) Registry: Method of construction. *Acta Geneticae Medicae et Gemellologiae*. 1987; 36:61–66. [PubMed: 3673478]
- Erickson RC, Scott ML. Clinical memory testing: a review. *Psychological Bulletin*. 1977; 84:1130–1149. [PubMed: 928573]
- Finkel D, Pedersen NL, McGue M. Genetic influences on memory performance in adulthood: Comparison of Minnesota and Swedish twin data. *Psychology and Aging*. 1995; 10:437–446. [PubMed: 8527064]
- Henderson WG, Eisen SE, Goldberg J, True WR, Barnes JE, Vitek M. The Vietnam Era Twin Registry: A resource for medical research. *Public Health Reports*. 1990; 105:368–373. [PubMed: 2116638]
- Jinks JL, Fulker DW. Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behavior. *Psychological Bulletin*. 1970; 73:311–349. [PubMed: 5528333]
- Johansson B, Whitfield K, Pedersen NL, Hofer SM, Ahern F, McClearn GE. 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: Series B: Psychological Sciences & Social Sciences*. 1999; 54B:P173–P179.
- Kendler KS, Heath AC, Martin NG, Eaves LJ. Symptoms of anxiety and depression in a volunteer twin population: The etiologic role of genetic and environmental factors. *Archives of General Psychiatry*. 1987; 43:213–221. [PubMed: 3954540]
- Kremen WS, Jacobsen KC, Xian H, Eisen SA, Eaves LJ, Tsuang MT, et al. Genetics of verbal working memory processes: a twin study of middle-aged men. *Neuropsychology*. 2007; 21(5):569–580. [PubMed: 17784805]
- Kremen WS, Jacobson KC, Panizzon MS, Xian H, Eaves LJ, Eisen SA, et al. Factor structure of planning and problem-solving: a behavioral genetic analysis of the Tower of London task in middle-aged twins. *Behavior Genetics*. 2009; 39:133–144. [PubMed: 19085096]
- Kremen WS, Thompson-Brenner H, Leung YJ, Grant MD, Franz CE, Eisen SA, et al. Genes, environment, and time: The Vietnam Era Twin Study of Aging (VETSA). *Twin Research and Human Genetics*. 2006; 9:1009–1022. [PubMed: 17254445]
- Kremen WS, Xian H, Jacobson KC, Eaves LJ, Franz CE, Panizzon MS, et al. Storage and executive components of working memory: integrating cognitive psychology and behavior genetics in the study of aging. *J Gerontol B Psychol Sci Soc Sci*. 2008; 63:P84–P91. [PubMed: 18441269]
- Luria, AR. *The working brain: An introduction to neuropsychology*. Haigh, B., translator. London: Allen Lane; 1973.
- Mariani E, Monastero R, Mecocci P. Mild cognitive impairment: A systematic review. *Journal of Alzheimer's Disease*. 2007; 12:23–35.
- McArdle JJ, Goldsmith HH. Alternative common factor models for multivariate biometric analyses. *Behavior Genetics*. 1990; 20:569–608. [PubMed: 2288547]
- McArdle JJ, Plassman BL. A biometric latent curve analysis of memory decline in older men of the NAS-NRC twin registry. *Behavior Genetics*. 2009; 39:472–495. [PubMed: 19404731]
- McClearn GE, Johansson B, Berg S, Pedersen NL, Ahern F, Petrill SA, et al. Substantial genetic influence on cognitive abilities in twins 80 or more years old. *Science*. 1997; 276:1560–1563. [PubMed: 9171059]
- Millis SR, Malina AC, Bowers DA, Ricker JH. Confirmatory factor analysis of the Wechsler Memory Scale-III. *Journal of Clinical and Experimental Neuropsychology*. 1999; 21:87–93. [PubMed: 10421004]
- Moore JH, Williams SM. New strategies for identifying gene-gene interactions in hypertension. *Annals of Medicine*. 2002; 34:88–95. [PubMed: 12108579]
- Neale, MC.; Boker, SM.; Xie, G.; Maes, HH. *Mx: Statistical Modeling*. 6th ed.. Richmond, VA: Department of Psychiatry, Medical College of Virginia; 2004.
- Neale, MC.; Cardon, LR. *Methodology for genetic studies of twins and families*. Dordrecht, The Netherlands: Kluwer Academic Publishers; 1992.

- Nee DE, Berman MG, Moore KS, Jonides J. Neuroscientific evidence about the distinction between short and long term memory. *Current Directions in Psychological Science*. 2008; 17:102–106.
- Nilsson LG. Memory function in normal aging. *Acta Neurologica Scandinavica*. 2003; 179:7–13.
- Papassotiropoulos A, Stephan DA, Huentelman MJ, Hoernndli FJ, Craig DW, Pearson JV, et al. Common Kibra alleles are associated with human memory performance. *Science*. 2006; 314:475–478. [PubMed: 17053149]
- Pedersen NL, Plomin R, Nesselroade JR, McClearn GE. A quantitative genetic analysis of cognitive abilities during the second half of the life span. *Psychological Science*. 1992; 3:346–353.
- Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. *Archives of Neurology*. 1999; 56:303–308. [PubMed: 10190820]
- Ranganath C, Blumenfeld RS. Doubts about double dissociations between short-and long-term memory. *Trends in Cognitive Sciences*. 2005; 9:374–380. [PubMed: 16002324]
- Reynolds CA, Finkel D, Gatz M, Pedersen NL. Sources of influence on rate of cognitive change over time in Swedish twins: An application of latent growth models. *Experimental Aging Research*. 2002; 28:407–433. [PubMed: 12227921]
- Schaie, KW. Intellectual development in adulthood: The Seattle Longitudinal Study. Cambridge: Cambridge University Press; 1996.
- Schaper K, Kolsch H, Popp J, Wagner M, Jessen F. KIBRA gene variants are associated with episodic memory in healthy elderly. *Neurobiology of Aging*. 2007
- Squire LR, Knowlton B, Musen G. The structure and organization of memory. *Annual Review of Psychology*. 1993; 44:453–495.
- Swan GE, Reed T, Jack LM, Miller BL, Markee T, Wolf PA, et al. Differential genetic influence for components of memory in aging adult twins. *Archives of Neurology*. 1999; 56:1127–1132. [PubMed: 10488814]
- Takashima A, Petersson KM, Rutters F, Tendolkar I, Jensen O, Zwarts MJ, et al. Declarative memory consolidation in humans: a prospective functional magnetic resonance imaging study. *Proceedings of the National Academy of Science*. 2006; 103:756–761.
- Thapar A, Petrill SA, Thompson LA. The heritability of memory in the Western Reserve Twin Project. *Behavior Genetics*. 1994; 24:155–160. [PubMed: 8024532]
- Tulving, E. *Elements of episodic memory*. Oxford, U.K.: Clarendon; 1983.
- Volk HE, McDermott KB, Roediger HL 3rd, Todd RD. Genetic influences on free and cued recall in long-term memory tasks. *Twin Research and Human Genetics*. 2006; 9:623–631. [PubMed: 17032541]
- Walhovd KB, Fjell AM, Dale AM, Fischl B, Quinn BT, Makris N, et al. Regional cortical thickness matters in recall after months more than minutes. *Neuroimage*. 2006; 31:1343–1351. [PubMed: 16540346]
- Wechsler, D. *Wechsler Abbreviated Scale of Intelligence*. San Antonio, Texas: The Psychological Corporation; 1999.
- Wiens AN, Tindall AG, Crossen JR. California Verbal Learning Test: A normative study. *The Clinical Neuropsychologist*. 1994; 8:75–90.
- Zec, RF. Neuropsychological function in Alzheimer's disease. In: Parks, RW.; Zec, RF.; Wilson, RS., editors. *Neuropsychology of Alzheimer's Disease and Other Dementias*. New York: Oxford University Press; 1993. p. 3-80.

Figure 1.
a. Cholesky Decomposition Model
b. Independent Pathway Model
c. Common Pathway Model

Figure 2. Standardized parameter estimates and variance components for the full independent pathway model

Parameters with an asterisk (*) are significantly greater than zero based on 95% confidence intervals. Values in parentheses are standardized variance components, calculated by squaring the individual parameter estimates, and represent the amount of variance in the variable of interest accounted for by each latent factor.

Figure 3. Standardized parameter estimates and variance components for the full common pathway model

Parameters with an asterisk (*) are significantly greater than zero based on 95% confidence intervals. Values in parentheses are standardized variance components, calculated by squaring the individual parameter estimates, and represent the amount of variance in the variable of interest accounted for by each latent factor.

Figure 4. Standardized parameter estimates and variance components for the best-fitting model (reduced common pathway model)

Parameters with an asterisk (*) are significantly greater than zero based on 95% confidence intervals. Values in parentheses are standardized variance components, calculated by squaring the individual parameter estimates, and represent the amount of variance in the variable of interest accounted for by each latent factor.

Table 1

Descriptive statistics and CVLT-2 variables stratified by zygosity

	MZ Twins (N=671)	DZ Twins (N=523)
Age (years)	55.3, 2.5	55.6, 2.4
Education (years)	13.8, 2.1	13.9, 2.2
Employment (% full time)	78.6%	77.2%
Ethnicity (% of sample)		
Caucasian	88.9%	91.6%
African American	4.8%	3.5%
Hispanic	3.2%	2.0%
CVLT Variables		
Trials 1–5 Total Score	43.4, 8.0	42.3, 9.1
Short Delay Free Recall	8.7, 2.8	8.6, 2.7
Long Delay Free Recall	9.1, 2.9	9.0, 2.9

With the exception of employment status and ethnicity, data are presented as means and standard deviations.

Table 2

Phenotypic, genetic, and environmental relationships between measures of immediate and delayed recall

	Trials 1–5 Total Score	Short Delay Free Recall	Long Delay Free Recall
Phenotypic			
Trials 1–5 Total Score	1.00		
Short Delay Free Recall	.75 (.72–.77)	1.00	
Long Delay Free Recall	.74 (.71–.76)	.83 (.81–.84)	1.00
Additive Genetic			
Trials 1–5 Total Score	.36 (.10–.55)		
Short Delay Free Recall	.89 (.55–1.0)	.24 (.05–.48)	
Long Delay Free Recall	.87 (.60–1.0)	.99 (.92–1.0)	.35 (.10–.56)
Shared Environment			
Trials 1–5 Total Score	.08 (<i>ns</i>)		
Short Delay Free Recall	.93 (<i>ns</i>)	.15 (<i>ns</i>)	
Long Delay Free Recall	.91 (<i>ns</i>)	.99 (<i>ns</i>)	.12 (<i>ns</i>)
Unique Environment			
Trials 1–5 Total Score	.56 (.48–.66)		
Short Delay Free Recall	.65 (.59–.71)	.61 (.53–.70)	
Long Delay Free Recall	.62 (.55–.68)	.71 (.66–.75)	.53 (.46–.62)

Standardized variance components are presented in *italics* on the diagonals of the additive genetic, shared and unique environment elements. Phenotypic, genetic and environmental correlations are presented on the off-diagonals. 95% confidence intervals are presented in the parentheses. Values that are not significantly different from zero based on 95% confidence intervals are noted by *ns*.

Table 3

Multivariate model fitting results

		-2LL	DF	LRT	ADF	p	AIC
1.	Saturated Model	7513.827	3528	-	-	-	-
2.	Trivariate Cholesky	7556.576	3561	42.749	33	.119	-23.251
3.	Genetic Independent Pathway Model	7556.840	3561	43.013	33	.114	-22.971
4.	Shared Environment Independent Pathway Model	7558.245	3561	44.418	33	.089	-21.582
5.	Unique Environment Independent Pathway Model	7556.576	3561	42.749	33	.119	-23.251
6.	Full Independent Pathway Model	7556.786	3561	42.959	33	.115	-23.041
7.	Full Common Pathway Model	7560.488	3565	46.661	37	.133	-27.339
8.	Reduced Common Pathway Model	7562.354	3570	48.527	42	.227	-35.473
9.	Measurement Model	7584.588	3571	70.761	43	.005	-15.239

-2LL = Negative 2 log-likelihood; DF = Degrees of freedom; LRT = Likelihood-ratio chi-square test; ADF = Change in degrees of freedom; AIC = Akaike's information criterion. The fit of models 2 through 9 are tested against that of model 1.