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

Boo Johansson,1,2 Keith Whitfield,2,4 Nancy L. Pedersen,3 Scott M. Hofer,4 Frank Ahern,2,4 and Gerald E. McClearn1,2,4

1Institute of Gerontology, University College of Health Sciences, Jönköping, Sweden.  2Department of Biobehavioral Health, Pennsylvania State University, University Park.  3Institute of Environmental Medicine, The Karolinska Institute, Stockholm, Sweden.  4Center for Developmental and Health Genetics, Pennsylvania State University, University Park.

The relative importance of genetic and environmental influences on episodic memory in very late life was studied using a quantitative genetic approach. Identical (n = 125) and same-sex fraternal (n = 157) twin pairs, aged 80 and older (mean age = 83.3; SD = 3.1) and without a diagnosis of dementia were tested with seven memory measures: (1-2) Digit Span Forward and Backwards, (3) Prose Recall, (4) Thurstone's picture memory test, and the Memory in Reality (MIR) test, including the subtasks of (5) free recall, (6) recognition, and (7) relocation. Heritabilities, estimated by structural equation modeling, ranged from .04 to .49. The digit span backward test showed the highest heritability (h² = .49), while heritabilities were typically lower for the long-term memory measures. The results demonstrate genetic influences on memory in the oldest-old, but suggest that the magnitude of these effects differs across memory measures.

MEMORY is one of the most studied aspects of cognition in late adulthood and aging (for reviews, see Craik & Salthouse, 1992; Kausler, 1994). Since the earliest psychometric studies of intelligence, memory has been a central construct in research on aging. However, it is typically studied as an element within a broader hierarchical structure of intellectual abilities (e.g., Horn & Hofer, 1992; Schaie, 1996). In experimental cognitive aging research, memory is recognized to be a highly complex, multidimensional construct in which various systems and processes are differentially affected by aging (Colscher, 1992; Craik & Jennings, 1992; Hultsch & Dixon, 1990; Richardson-Klavehn & Bjork, 1988). A general finding is that free recall of episodic memories is the most affected memory operation in normal aging, whereas retrieval of semantic and procedural memories remains fairly stable. Issues of emerging importance in aging research include assessing interindividual differences in memory performance, and the influence of age by studying very old samples.

Interest in the oldest-old arises from the increasing number of individuals over 80 years of age in industrialized countries (Kinsella & Tauber, 1993) and incidence rates of dementia (e.g., Johansson & Zarit, 1995) in the oldest-old make it obligatory to adjust for dementia because of its deleterious effects, especially on episodic memory (Bondi et al., 1994; Sliwinski et al., 1996).

Variability across individuals in memory performance is a well-observed phenomenon. Interindividual variation in memory may differ not only as a function of age and dementia, but also in respect to the type of memory under study (for reviews see Craik & Jennings, 1992; Hultsch et al., 1991; Richardson-Klavehn & Bjork, 1988). Because there are numerous dimensions for various components of memory which can be assessed using different methodological strategies (Richardson-Klavehn & Bjork, 1988), evaluation of individual differences necessarily requires a battery of memory measures.

Interindividual differences represent the basis for quantitative genetics, as its analytical approach is based on the decomposition of variance in performance into the most fundamental sources—genetic and environmental components (e.g., Falconer, 1981; Flomir et al., 1997). Quantitative genetic research, however, requires genetically informative samples of individuals that are related in varying degrees. A common design is the "classic twin design," which compares the similarity of identical and fraternal twins.

There are, so far, only a few quantitative genetic studies of memory. For example, Pedersen and colleagues (1992) report heritabilities for episodic memory tasks in the range of 32% to 44% in the Swedish Adoption/Twin Study of Aging (SATSA; average age 66). Finkel and McGue (1993) investigated episodic tasks such as word recall, text recall, and figure memory in the Minnesota Twin Study of Adult Development and Aging (MTSADA; mean age 67) and found somewhat higher heritabilities (between .50 and .60). In further analysis, Finkel, Pedersen,
McGue, and McClearn (1995) compared several cognitive tests
(including span memory) across these two samples and found
little evidence for age differences in heritability, with heritability
estimates of span memory around .30 (independent of its rela-
tively low contribution to a general ability factor). In a meta-
analysis, Thapar, Petrill, Whitfield and Thompson (unpublished
manuscript), also showed that there are significant differences in
heritabilities for visual and verbal forms of memory. In a recent
article, we examined memory and other cognitive abilities at the
factor level in the sample of the oldest-old used for the present
study. Memory was defined as a composite of the scores in the
Digit Span and Thurstone’s picture memory tests (see Measures
later). The heritability for the memory composite was estimated
to be 52% (McClearn et al., 1997), which can be compared to
39% and 64% in the older subgroups in the SATSA and
MTSADA samples, respectively.

Sources of individual differences in episodic memory mea-
sures in very late life are not well understood. For example, the
evidence is somewhat mixed as to whether genetic influences
remain at constant levels, are magnified, or lose their importance
for measures indicating aspects of short- and long-term memory
performance. The purpose of this study was to examine the rela-
tive importance of genetic and environmental influences on a va-
riety of memory measures in a sample of the oldest-old.

METHODS

Sample

Potential candidates for participation in this study were 351
twin pairs (149 identical or monozygotic, MZ, twin pairs, and
202 same-sex fraternal or dizygotic, DZ, twin pairs) aged 80 and
older, who were assessed in the first wave of the ongoing lon-
gitudinal study “Origins of Variance in the Old-Old” (OCTO-Twin
Study; Johansson & McClearn, 1996). The sample was drawn
from the oldest-cohort of the population-based Swedish Twin
Registry (Cederlöf & Lorich, 1978), which was comprised of all
complete twin pairs, born 1913 and earlier, who were both alive
when contacted for potential participation if they were, or be-
came, 80 years of age during the three-year period of data col-
lection that started in 1991 (737 pairs in 1474 individuals). Of
these pairs, some were excluded because one or both were de-
ceased before they were scheduled for examination (188 pairs),
or because one or both declined participation in the study for
other reasons (198 pairs). Seven hundred and two individuals in
351 complete twin pairs participated. Other than for reasons of
death, the pairwise cooperation rate at the initiation of this study
was 65% (corresponding to an individual response rate of 80%).
A comparison between participants and nonparticipants showed
no differences for age, gender, or type of housing (Johansson &
McClearn, 1996). The gender ratio, education, socioeconomic
status, marital status, and housing in the investigated sample cor-
responds to population statistics for this age segment of the
Swedish population (Simmons et al., 1997). In addition, there
were no significant differences between MZ and DZ twins for
age, marital status, or type of housing.

For the present analyses, 69 pairs (81 individuals) were ex-
cluded because one or both in the pair met the DSM-III-R cli-
cal criteria for dementia (APA, 1987). The nondemented sample
(n = 564 individuals) for the present analyses comprised 125
pairs of identical and 157 pairs of fraternal twins (see Appendix,
Note 1). Their mean age was 83.3 (median = 82.3) with a range
from 79.4 to 97.9 years. Sixty-four percent were women. For fur-
ther details see Table 1.

Measures

The OCTO-Twin Study includes a broad spectrum of biobe-
havioral measures of health and functional capacity, personal-
ity, well-being, and interpersonal functioning, as well as mem-
ory and cognition. Cognitive functioning is assessed with a
widely used Swedish psychometric battery (SRB; Dureman &
Sälde, 1959) comprising tests for verbal meaning/synonyms,
inductive logical reasoning, spatial ability, and perceptual
speed. In addition, a brief battery of neuropsychological tests
was administered, including tests of orientation, calculation
(The Coin Test), spatial ability (The Clock Test), and the Mini-
Mental State Examination Test (Folstein et al., 1975).

The memory battery assesses aspects of short-term and long-
term memory systems, involving episodic memory processing
of verbal and nonverbal material, with different demands on the
retrieval mechanisms (recall and recognition). The following
tests were used in the present study:

(1-2). The Digit Span Test measures short-term memory for
orally presented digits (Weschler, 1991). In the forward
part (1), subjects are asked to recall the digits in the same
order as they were presented, whereas in the backward
part (2) they are instructed to recall the digits in reverse
order. The maximum score is 9 for the forward, and 8
for the backward part of the test.

(3). The Prose Recall Test is a verbal memory test where
subjects are asked for immediate free recall of a brief
story (100 words) that has a humorous point (Johansson
et al., 1992). Responses are coded for the amount of infor-
mation recalled in a manner similar to the Wechsler
Memory Test (WMS; Wechsler, 1945). The maximum
score is 16.

Table 1. Descriptive Statistics for the Twin Sample

<table>
<thead>
<tr>
<th></th>
<th>MZ</th>
<th>DZ</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N = 230)</td>
<td>(N = 314)</td>
<td>(N = 564)</td>
</tr>
<tr>
<td>Age (Mean years)</td>
<td>83.5 (3.2)</td>
<td>83.2 (3.0)</td>
<td>83.3 (3.1)</td>
</tr>
<tr>
<td>Range</td>
<td>79.4-97.9</td>
<td>79.7-97.8</td>
<td></td>
</tr>
<tr>
<td>Gender (No. males/females)</td>
<td>106/144</td>
<td>98/216</td>
<td>204/360</td>
</tr>
<tr>
<td>% males</td>
<td>42.4</td>
<td>31.2</td>
<td>36.2</td>
</tr>
<tr>
<td>Education (Mean years (SD))</td>
<td>7.6 (2.8)</td>
<td>6.9 (2.0)</td>
<td>7.2 (2.4)</td>
</tr>
<tr>
<td>Marital status % married</td>
<td>32.5</td>
<td>32.8</td>
<td>32.7</td>
</tr>
<tr>
<td>% widowed</td>
<td>53.4</td>
<td>51.6</td>
<td>52.4</td>
</tr>
<tr>
<td>% never married/divorced</td>
<td>14.0</td>
<td>15.6</td>
<td>14.9</td>
</tr>
<tr>
<td>Housing % ordinary</td>
<td>91.1</td>
<td>87.9</td>
<td>88.8</td>
</tr>
<tr>
<td>% service housing</td>
<td>6.5</td>
<td>9.8</td>
<td>8.3</td>
</tr>
<tr>
<td>% institution</td>
<td>2.4</td>
<td>3.3</td>
<td>2.9</td>
</tr>
</tbody>
</table>

Note. Standard deviation is shown in parentheses.
(4). *Thurstone's picture memory* test is a nonverbal, long-term memory test (Thurstone & Thurstone, 1949). Subjects are shown 28 pictures and then asked for recognition of these among others. The pictures were enlarged from the original version to minimize any possible visual problems. The maximum score is 28.

(5–7). The *Memory-in-Reality (MIR)* Test first requires the naming of 10 common real-life objects shown to the subject. The subjects are then instructed to place these objects in the different rooms of a three-dimensional model of an apartment, according to their own preferences. Thirty minutes later they are asked for free recall of the objects (5) followed by a recognition task (6) for the objects that were not recalled. Subjects are then asked to place the objects in the same locations as they did previously—the relocation test (7). The maximum score in each subtest is 10. The recognition score is the sum of free recall and the items recognized (Johansson, 1988; 1989).

Test-retest and split-half reliability for Digit Span are typically found in the range 72–83. The split-half reliability coefficient for Prose Recall is .88 (see Johansson et al., 1997), and .82 for Thurstone's Picture Memory Test (see Berg, 1980). For the MIR tasks, the split-half reliabilities were .88 for free recall, .96 for recognition, and .90 for relocation. The reliability coefficients across tests are all in a range that allows variance decomposition at the test levels for the sake of estimating the relative contributions of environmental and genetic influences. (For more details about the test battery and its validity see Johansson et al., 1992; Johansson & Zarit, 1995).

**Procedures**

The twins are investigated in their place of residence by licensed and experienced nurses (RNs), specially trained for the study, and continuously supervised. A complete testing session lasts approximately 3.5–4.0 hours and includes rest periods. Scheduling was arranged to minimize geographical, age, or gender order effects.

Both members in a pair were investigated by a different nurse within one month of the cotwin's test date. The nurses were deliberately kept blind to zygosity of the twins to avoid expectation biases. In pairs where zygosity was unknown, a diagnosis was based on DNA fingerprinting (Johansson & McClearn, 1996).

### Table 2. Means and Standard Deviations for the Memory Tests

<table>
<thead>
<tr>
<th>Test</th>
<th>MZ</th>
<th>DZ</th>
<th>Total</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digit Span</td>
<td>5.6 (1.2)</td>
<td>5.3 (1.3)</td>
<td>5.4 (1.3)</td>
<td>530</td>
</tr>
<tr>
<td>forward</td>
<td>3.4 (1.5)</td>
<td>3.2 (1.6)</td>
<td>3.3 (1.5)</td>
<td>529</td>
</tr>
<tr>
<td>backward</td>
<td>9.9 (4.0)</td>
<td>9.3 (4.1)</td>
<td>9.5 (4.1)</td>
<td>483</td>
</tr>
<tr>
<td>Prose Recall</td>
<td>18.7 (4.9)</td>
<td>18.0 (5.4)</td>
<td>18.3 (5.2)</td>
<td>388</td>
</tr>
<tr>
<td>Picture Memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIR</td>
<td>6.4 (2.5)</td>
<td>6.2 (2.5)</td>
<td>6.3 (2.5)</td>
<td>486</td>
</tr>
<tr>
<td>free recall</td>
<td>9.5 (1.6)</td>
<td>9.4 (1.7)</td>
<td>9.5 (1.7)</td>
<td>487</td>
</tr>
<tr>
<td>recognition</td>
<td>7.8 (2.5)</td>
<td>7.6 (2.4)</td>
<td>7.7 (2.4)</td>
<td>481</td>
</tr>
</tbody>
</table>

*Note. Standard deviations are shown in parentheses. Estimates based on all available cases (MZn = 173–236; DZn = 215–294).*

### Missing Data

Within-occasion missing data on the memory measures were due to vision and hearing difficulties, as well as fatigue over the testing session. Approximately half of the participants had complete data on all memory subtests: 47.2% of MZ twin pairs (N = 59) and 45.2% of DZ twin pairs (N = 71). A substantial number of individuals within twin pairs were missing only a single memory test (95 MZ twin pairs [76%]; 107 DZ twin pairs [68%]). Few individuals were missing all seven memory measures (4 MZ and 7 DZ cotwins).

### Statistical Analysis

Descriptive statistics and phenotypic intercorrelations among the memory measures were calculated first. A principal component analysis was performed to examine the structure of the correlations among the memory measures. Intraclass correlations were then computed (controlling for the effects of Age and Gender) separately for MZ and DZ pairs. These intraclass correlations were computed based on available pairwise cases and for nondemented pairs using maximum likelihood estimation.

Structural equation modeling using Mx (Neale & Cardon, 1992; Neale, 1995) was used to perform variance component analysis of the genetic and environmental influences on the memory measures. Two covariates, Age and Gender, were included in the model to regress out the potential confounding effects from the variance components. This analysis was performed on all individuals regardless of missing information within-person or twin dyad. Unbiased estimates of population parameters were obtained using the direct maximum likelihood approach under the assumption that the data are missing at random (Graham, Hofer, Donaldson, MacKinnon, & Schafer, 1997; Little & Rubin, 1987).

### Results

**Descriptives: Zygosity, Age, and Gender**

Inspection of the distributional properties of the memory measures demonstrated minimal skewness and kurtosis with the exception of MIR Recognition. Over 85% of individuals scored 10 out of 10 on MIR Recognition with the remaining individual scores distributed across the lower range of the measure. This measure was dropped from subsequent analyses. For the other memory measures, few individuals had zero values on three or more (out of seven) memory measures (within 1.6% MZ and 2.5% DZ twin pairs).

A series of independent-groups t tests were performed to test for differences in memory performance between zygosity and gender groups. No significant differences in performance were found between MZ and DZ twins, shown in Table 2. Significant gender effects (regardless of zygosity) were found for Digit Span forward, t(528) = 2.1; p < .05 (higher performance in men) and MIR Recall and Relocation, t(484) = -3.4; p < .01, and t(479) = -3.7; p < .01, respectively (higher performance in women). In addition, significant correlations with age (lower performance with age) were found for all of the memory measures (correlations shown in Table 3).

**Phenotypic correlations among memory measures**

A principal component analysis, with oblique rotation (pro-max rotation to a varimax target matrix), was performed to ex-
amine the structure of the correlations between the memory measures (shown in Table 3). This analysis was performed using only cases with complete data (n = 358) for all memory measures, regardless of zygosity, and treating each member of a twin dyad as a single case.

Examination of a scree plot indicated two components. A two-component solution accounted for 72.2% of total variance with off-component loadings less than .15. The first component accounted for 34% of variance (independent of the second component) and was identified by measures indicating long-term memory (Prose recall [.69], Picture memory [.81] and the MIR Recall [.62] and Relocation [.71] subtests). The second component was clearly indicated by the two span memory measures, Digit Span Forward (.82) and Backwards (.76), and accounted for approximately 22% variance independent of the first component. The two components were moderately correlated (r = .46).

**Intraclass Correlations**

Intraclass correlations, presented in Table 4, were calculated controlling for differences associated with age and gender (McGue & Bouchard, 1984). Under various assumptions (e.g., equivalent environments), the presence of genetic influences is indicated when the intraclass correlations for MZ pairs are greater than those for DZ pairs. This is the case in all of the measures with the exception of Prose Recall, where the intraclass correlations are equivalent.

**Variance Components Analysis**

Structural equation modeling using Mx (Neale, 1995) was employed to estimate genetic, shared environmental, and non-shared environmental components of variance. The modeling approach allows the examination of how well the expectations of certain genetic models fit the observed MZ and DZ (intra-class) covariance structure. Because the expectation for genetic correlations between MZ and DZ cotwins is 1.00 and .50, respectively, the correlation between MZ cotwins is fixed at 1.00. The genetic correlation between DZ cotwins is set at .50 in an additive model (Neale & Cardon, 1992). The variance not accounted for by the genetic expectations is ascribed to environmental influences, shared and nonshared (unique environmental and error variance). Because the OCTO-Twin design is based on twins with the same rearing conditions, a full separation of shared environmental and nonadditive (i.e., dominance effects and epistasis) sources of variance is not possible. All subsequent analyses were therefore aimed at the decomposition of the total variance into additive genetic, shared environmental, and nonshared environmental sources.

The memory measures were regressed on Age and Gender simultaneously in order to remove potential confounding variance (since the memory measures were age-related and some gender differences were indicated). The estimates of genetic and environmental sources of variance were computed based on the total phenotypic variance after statistically controlling for variance associated with Age and Gender. We report estimates on the full nondemented sample (with partially complete data) only.

The proportion of total variation due to genetic variance (see Table 5) for Digit Span was 27% in the forward part and 49% in the backward part of the test; 4% for Prose Recall, 47% for Picture Memory, 28% for MIR recall, and 37% for MIR Relocation.

**Table 3. Phenotypic Correlations for the Memory Measures and Age**

<table>
<thead>
<tr>
<th>Variables</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0. Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Digits forward</td>
<td>-.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Digits backward</td>
<td>-.15</td>
<td>.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Prose Recall</td>
<td>-.20</td>
<td>.29</td>
<td>.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Picture memory</td>
<td>-.15</td>
<td>.22</td>
<td>.29</td>
<td>.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. MIR recall</td>
<td>-.21</td>
<td>.23</td>
<td>.32</td>
<td>.51</td>
<td>.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. MIR recognition</td>
<td>-.16</td>
<td>.25</td>
<td>.28</td>
<td>.37</td>
<td>.45</td>
<td>.50</td>
<td></td>
</tr>
<tr>
<td>7. MIR relocation</td>
<td>-.19</td>
<td>.20</td>
<td>.31</td>
<td>.49</td>
<td>.52</td>
<td>.62</td>
<td>.42</td>
</tr>
</tbody>
</table>

*Note. All correlations are significant (p < .001). Estimates are based on all available pairwise cases with the minimum N = 384.*

**Table 4. Intraclass Correlations for the Memory Measures**

<table>
<thead>
<tr>
<th></th>
<th>Pairwise MZ</th>
<th></th>
<th></th>
<th>Maximum Likelihood MZ</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(incomplete data)</td>
<td></td>
</tr>
<tr>
<td>1. Digits forward</td>
<td>.38** (113)</td>
<td>.32** (139)</td>
<td>.40**</td>
<td>.30**</td>
<td></td>
</tr>
<tr>
<td>2. Digits backward</td>
<td>.49** (113)</td>
<td>.22* (138)</td>
<td>.50**</td>
<td>.20**</td>
<td></td>
</tr>
<tr>
<td>3. Prose Recall</td>
<td>.42** (98)</td>
<td>.44** (112)</td>
<td>.45**</td>
<td>.46**</td>
<td></td>
</tr>
<tr>
<td>4. Picture memory</td>
<td>.33** (65)</td>
<td>.25* (84)</td>
<td>.39**</td>
<td>.20**</td>
<td></td>
</tr>
<tr>
<td>5. MIR recall</td>
<td>.29** (99)</td>
<td>.09 (115)</td>
<td>.33**</td>
<td>.11</td>
<td></td>
</tr>
<tr>
<td>6. MIR recognition</td>
<td>——</td>
<td>——</td>
<td>——</td>
<td>——</td>
<td></td>
</tr>
<tr>
<td>7. MIR relocation</td>
<td>.39** (98)</td>
<td>.12 (112)</td>
<td>.43**</td>
<td>.13</td>
<td></td>
</tr>
</tbody>
</table>

*Note. All intraclass correlations were partialled for Age and Gender. Sample sizes for pairwise available twin pairs are shown in parentheses. Maximum likelihood (ML) estimates were based on the full nondemented sample with incomplete data (MZn = 125; DZn = 157). Statistically significant intraclass correlations are shown by asterisks (** = p < .01, * = p < .05).*

**Table 5. Estimates of Variance Components for the Memory Measures**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Parameter estimates</th>
<th>Genetic A</th>
<th>SE</th>
<th>Environment NSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Digits forward</td>
<td>.27 (.00,.56)</td>
<td>.16 (.00,.43)</td>
<td>.57 (.44,.74)</td>
<td></td>
</tr>
<tr>
<td>2. Digits backward</td>
<td>.49 (.18,.60)</td>
<td>.00 (.00,.22)</td>
<td>.51 (.40,.66)</td>
<td></td>
</tr>
<tr>
<td>3. Prose Recall</td>
<td>.04 (.00,.46)</td>
<td>.43 (.09,.56)</td>
<td>.53 (.39,.66)</td>
<td></td>
</tr>
<tr>
<td>4. Picture memory</td>
<td>.47 (.00,.63)</td>
<td>.00 (.00,.36)</td>
<td>.53 (.37,80)</td>
<td></td>
</tr>
<tr>
<td>5. MIR recall</td>
<td>.28 (.00,.44)</td>
<td>.00 (.00,.26)</td>
<td>.72 (.56,90)</td>
<td></td>
</tr>
<tr>
<td>6. MIR recognition</td>
<td>——</td>
<td>——</td>
<td>——</td>
<td></td>
</tr>
<tr>
<td>7. MIR relocation</td>
<td>.37 (.01,.52)</td>
<td>.00 (.00,.27)</td>
<td>.63 (.48,80)</td>
<td></td>
</tr>
</tbody>
</table>

*Note. Parameter estimates (95% confidence intervals shown in parentheses) are based on full nondemented sample with incomplete data (MZn = 125; DZn = 157). A refers to additive genetic influences, SE refers to shared environment, and NSE to nonshared environmental influences. *indicates that no model fitting was conducted for this measure. Variance components are based on total phenotypic variance after controlling for variance associated with both age and gender.*
Although several of the confidence intervals around the genetic variance estimates include zero, these results are consistent with other studies showing genetic influence on memory performance. The confidence intervals of two of the measures, Digit Span Backward and MIR relocation, indicate statistical significance.

Inspection of shared environmental sources or familial influences revealed a significant effect for Prose Recall, which accounted for 43% variance. Shared environmental variance for Digit Span Forward was estimated to be 16%, but was not statistically significant. Nonshared (unique) environmental influences ranged from .51 (Digit Span Backwards) to .72 (MIR Recall).

DISCUSSION
The focus of this study was the relative influences of genetic and environmental sources of individual differences in specific tests of episodic memory functioning in nondemented individuals in very late life.

The most pronounced finding was that more than 50% of the variance associated with memory performance could be ascribed to nonshared environmental influences. Some of this nonshared influence is due to unreliable sources of variance, rather than systematic variability unique to the individual. The estimates that we present must, indeed, be considered as lower-bound estimates, because removing the unreliable proportions of variance would increase the relative proportions of genetic and shared-environmental sources of variability. It is also the case that comparisons of proportions of variance across memory measures must be made with caution if there is evidence for differences in reliability across measures of various aspects of memory performance. In the following discussion we assume that the reliability of the employed measures is sufficiently equal to allow comparisons across measures. The broad confidence intervals associated with the estimates, however, emphasize the need to interpret the findings with caution.

The results support previous findings that the magnitude of genetic influences on memory varies across tests (Thapar et al., 1994), although the sample size of our study is not sufficient to conclude that these estimates were statistically significant. For example, while we found a substantial heritability (49%) for the Digit Span backward task, a much lower genetic influence was found for the Forward Span (27%) of this short-term memory test. We suggest that the differential heritability pattern in this case reflects different complexities in performing the two tasks. Forward Span requires no elaboration of the to-be-remembered material, mostly attention and concentration necessary for temporal storage. More cognitive effort is required in the backward part where scanning and recall of the digits has to be performed in a reverse order, compared to the encoding for temporal short-term storage. The higher heritability in the backward task may reflect the involvement of more complex cognitive processes. Parallel support for this tentative hypothesis is provided by findings of greater heritability for other types of cognitive abilities, such as reasoning and verbal and spatial abilities (e.g., McClearn et al., 1997).

The results from the indicators of long-term memory, Prose Recall, Picture Memory, and the subtasks in the MIR test, also present an interesting finding in the context of heritability and cognitive load. For example, the heritability estimates for Prose Recall and MIR recall differed (4% and 28%, respectively), although both tests are verbal. These tests also demonstrate a large nonshared environmental component. In addition, Prose Recall showed a substantial shared environment component that may suggest the importance of language and verbal facilitation in early life, and the maintenance of such environments in later life. The MIR and Prose tests vary considerably in their respective loads at encoding, as well as in subsequent access to retrieval cues. Prose recall requires sustained attention at the experimenter-defined, one-time presentation (the 100-word story is orally presented at a certain standard pace); the 10 MIR objects are presented in a nonspeeded format that requires subjects to handle the objects when placing them in different rooms of the apartment model. The story's meaning and humorous point may serve to facilitate retrieval processes in Prose recall, whereas the individual can use imagery at MIR recall. In both tests subjects are provided with contextual cues that lower the cognitive expenditure and facilitate performance. The different magnitudes of genetic influences may then reflect differences in the cognitive processes required in performing these episodic tests.

Compared to the recall tests, featuring a two-process operation involving both search of information and decision, Picture Memory only involves the latter, as cues are physically present at testing. The cognitive expenditure is also affected by dual-coding (e.g., Hasher & Zacks, 1979) and more automatic processing of pictures (Hasher & Zacks, 1979). In terms of a cognitive load hypothesis, which would suggest lower genetic influences for less cognitively demanding tasks, the Picture Memory test exhibited a moderate heritability (47%).

Interestingly, the next highest heritability estimate among the long-term memory measures was found for the MIR-relocation task (37%). Although it is a subject-performed-task (e.g., Bäckman & Nilsson, 1985) featuring facilitation by multimodality at encoding of the objects, and reminding at prior recall and recognition tests, there was no dramatic performance improvement in this memory for location task or relocation when compared with free recall (see Table 2; Johansson, 1985). The relocation test, however, involves an additional spatial memory ("where did I put this object?") and strategy-dependency component (the appropriateness of how objects initially were placed).

There were four measures not used previously in quantitative genetic studies (MIR relocation, relocation, and the Prose Recall Test). Prose Recall showed a large, nonshared environmental component and lower heritability than a comparable test of Test Memory performance in late adulthood (Finkel et al., 1995). For the Picture Memory test, our heritability estimate (47%), was in the same range as estimates reported previously for young (39%), middle-aged (50%), and older (30%) groups in the SATSA study (Finkel et al., 1995). These cross-sectional comparisons suggest stability in heritability with age, but require confirmation in longitudinal studies. The finding of genetic influences in Digit Span Forward and Backwards tests corresponds with results from the SATSA sample of predominantly young-old twins (Pedersen et al., 1992).

The results support and extend previous findings from somewhat younger samples that genetic sources of variance appear also for memory performances in late life (e.g., Finkel et al., 1995). In this sense, our findings also begin to address the issue of whether heritability continues to increase, remains similar into late adulthood, or decreases at the end of the life span. The prevailing assumption in nongenetic theories of aging is that environmental influences necessarily accumulate throughout the life span.
span, which would predict lower heritabilities in the oldest-old when compared with earlier in life, because genetic influences would presumably remain stable, while total phenotypic variability would increase. In a meta-study of cognitive abilities across the life span, Pedersen and Lichtenstein (1997) show that the current literature seems to provide tentative support for this hypothesis. The data on memory functioning, however, are still too sparse to provide evidence for a comparable developmental scenario.

In the gerontological literature, variation in memory performance in aging has been attributed to interindividual differences in speed of processing (e.g., Park et al., 1996; Saltounse, 1985), sensory functioning (Lindenberger & Baltes, 1994), physical and mental health (e.g., Bäckman et al., 1996; Kalra et al., 1994; Nyberg et al., 1996; Wahlin et al., 1996), and dementia (e.g., Silviswski et al., 1996). These influences tend to become more evident with age. In subsequent multivariate modeling there is a need to include this broader assay of variables in the analyses (see Bergeman, 1997; Whitfield, 1994). The inclusion of such candidate variables can provide better insight into the origins of biomedical, as well as psychosocial, factors that influence memory in aging. In this context, the present study represents the first step to identify the relative effects of environmental and genetic influences on specific memory processes.

Our results suggest that heritability estimates on memory tests reflect the genetic influences on mixtures of cognitive processes required to perform a certain task. An inclusion of a broader memory battery elaborating on cognitive demands and encompassing measures of semantic and procedural memory, besides short-term memory and episodic long-term memory tests, is therefore suggested for future quantitative genetic studies of cognition. We would also encourage the use of multiple indicators of various aspects of memory performance which would allow a focus on reliable sources of individual differences.

The finding of a substantial heritability for the Digit Span Backwards suggests that this measure may be a valid “memory candidate” for molecular techniques, such as quantitative trait loci (see Plomin et al., 1997). Quantitative genetic studies like OCTO-Twin provide valuable information to guide molecular genetic studies for not only the investigation of memory (e.g., Nilsson et al., 1996), but for other phenotypes of interest to gerontological research (Plomin et al., 1997).

In conclusion, the results from this quantitative genetic study of memory in the oldest-old demonstrate a need for further analyses at the level of specific types of memory processes. The genetic influences and/or shared environmental sources of variance were found to be low to moderate in this study and, in conjunction with findings from previous studies, show little evidence that these influences disappear in very old life.

ACKNOWLEDGMENTS

The OCTO-Twin Study ("The Origins of Variance in the Old-Old: Oecogenarian Twins") is an ongoing longitudinal study conducted at the Institute of Gerontology (IFG), University of Health Sciences in Jönköping, Sweden, in collaboration with the Center for Developmental and Health Genetics at the Pennsylvania State University (PSU), USA, and the Division of Epidemiology at the Karolinska Institute in Stockholm, Sweden.

The authors are also greatly indebted to the RN's Lena Ahlbäck (1991–), Agneta Carlholt (1993–), Gunilla Hjalmarsson (1991–1993), Eva Georgsson (1993–1995), and Anna Lena Wetterholm (1993–1994) who travelled throughout the country and examined the twins. Invaluable help in coordinating the study is provided by Inger Cronholm at IFG and Elna Pyle at PSU. The authors would also like to thank the reviewers for valuable comments. The study is supported by a grant from the National Institute on Aging (NIA: AG 08861) of the National Institutes of Health.

Address correspondence to Boo Johansson, Institute of Gerontology, University College of Health Sciences, SE - 551 11 Jönköping, Sweden. E-mail: boo.johansson@hhj.hj.se

REFERENCES

Johansson, B. (1985). Memory and memory measurement in old age: Memory structure, context and metamemory. Departments of Applied Psychology and Geriatric and Long-Term Care Medicine, University of Gothenburg and the Institute of Gerontology, Jönköping, Sweden (Diss.).
HERITABILITY OF MEMORY

P179


Received October 7, 1997

Accepted December 3, 1998

Appendix

Note 1. Numbers are based on information from the completed diagnostic work-up performed on all dementia suspects. Pairs excluded in the present analyses were those in which one or both members met the DSM-III-R criteria for dementia. The figures differ slightly from the sample analyzed in McClearn and colleagues (1997), where all pairs with dementia suspects were excluded.