Journal of Gerontology: PSYCHOLOGICAL SCIENCES
2006, Vol. 61B, No. 4, P245–P249

Copyright 2006 by The Gerontological Society of America

Stability of Sex Differences in Cognition in Advanced Old Age: The Role of Education and Attrition

Denis Gerstorf,[1] Agneta Herlitz,[1,2] and Jacqui Smith[1]

[1]Max Planck Institute for Human Development, Berlin, Germany.
[2]Aging Research Centre, Karolinska Institute, Stockholm, Sweden.

We examined whether patterns of sex differences on tasks of perceptual speed, episodic memory, verbal fluency, and verbal knowledge are maintained during advanced old age. Using incomplete 13-year longitudinal data from participants in the Berlin Aging Study screened for dementia (*N* = 368; *M* = 83 years; range 70–100 years at baseline assessment), we estimated sex-specific age trajectories of cognitive change and explored the contributing role of education and attrition. We found that women and men declined virtually in parallel, with no evidence of differential change. After we controlled for age cohort-related differences in education, women outperformed men on tasks in the four cognitive domains. Findings also provide initial evidence that sex differences might be masked by differential patterns of sample attrition.

CROSS-SECTIONAL studies have documented that men outperform women on tasks requiring visuospatial processing whereas women show an advantage on many episodic memory tasks and a variety of speeded tasks, including Digit Symbol Substitution, Finding As, and Picture Identification (e.g., Maitland, Herlitz, Nyberg, Bäckman, & Nilsson, 2004; Meinz & Salthouse, 1998). It is not clear whether these sex differences change in pattern and magnitude during old age and if they generalize to the oldest-old population (e.g., individuals 85 years of age or older; see Hassing, Wahlin, & Bäckman, 1998). Some longitudinal studies report preserved sex differences and parallel patterns of decline (e.g., Finkel, Reynolds, McArdle, Gatz, & Pedersen, 2003; Schaie, 2005). Others argue that women are especially vulnerable to decline in memory (e.g., Zelinski & Stewart, 1998).

Two sets of intertwined factors may influence sample-specific findings about sex differences in level and slope of cognitive change in advanced old age. First, there are effects attributable to gender- and cohort-linked inequalities in societal opportunity structures, such as access to formal and professional education, that have consequences for the development of cognitive reserve capacity (e.g., Elias, Elias, D'Agostino, Silbershatz, & Wolf, 1997; Seeman et al., 2005; Smith & Baltes, 1998). Second, there are sex differences in dementia incidence and longevity that may play a role in sample attrition (e.g., Fratiglioni et al., 1997; Rabbitt, Lunn, & Wong, 2005).

We applied multilevel modeling to estimate sex-specific age trajectories in normal cognitive aging on eight tasks for a subsample of incomplete 13-year longitudinal data from the Berlin Aging Study (BASE; Baltes & Mayer, 1999). The present study thus extends previous cross-sectional analyses (Lindenberger & Baltes, 1997) and 6-year longitudinal analyses of sex differences in cognition in the BASE (see Singer, Verhaeghen, Ghisletta, Lindenberger, & Baltes, 2003). Unlike researchers in earlier BASE reports, we specifically exclude participants with suspected dementia and examine education and sex-specific attrition effects. We also evaluate sex differences at the level of individual cognitive tests rather than ability composites, because the effects of sex may differ by test within an ability and be masked by aggregating across tests.

METHODS

Participants and Procedure

We analyzed a subsample of BASE data from five measurement occasions collected over 13 years: Time 1 (T1; 1990–1993), Time 3 (T3; 1995–1996), Time 4 (T4; 1997–1998), Time 5 (T5; 2000), and Time 6 (T6; 2004–2005). (An additional assessment in 1993–1994 [Time 2] involved only one session and could not be included here because it did not provide comprehensive cognitive data.) Table 1 provides sample and descriptive information. T3 took place 3.76 years after T1 ($SD = 0.76$); T4, 5.48 years ($SD = 0.80$); T5, 8.94 years ($SD = 0.87$); and T6, 12.99 years ($SD = 0.90$) after T1, respectively. Comprehensive information about the T1, T3, and T4 samples and cognitive measures is published in Baltes and Mayer (1999) and Lövdén, Ghisletta, and Lindenberger (2004). We excluded from our analyses those participants at each occasion who were below an age cohort-specific cutoff on a dementia screening instrument (Short Mini-Mental State Examination, or SMMSE; Klein et al., 1985; 70–84 years, <12 points; 85+ years, <11 points). Independent clinical diagnoses of dementia in the BASE at T1 and T3 indicated sufficient specificity (72–98%) and sensitivity of the exclusion criteria (63–88%), both among the old-old and oldest-old women and men. Overall, we excluded 82 women and 66 men (see endnote 1). The average age of the total (incomplete) sample ($N = 751$) was 83.65 years ($SD = 7.20$; born between 1887 and 1922). Although each participant idiographically contributed different longitudinal time segments (observation period, $M = 3.4$ years; $SD = 4.6$; range = 0–15 years) to an age gradient spanning 30 years, the age information entered into the models was spread relatively equally across the three age decades (70s, $n = 237$; 80s, $n = 341$; 90s, $n = 173$).

Longitudinal sample attrition was primarily due to mortality: At T6, 83% of the participants in the T1 baseline sample were

Table 1. Descriptive Statistics for the Longitudinal Samples Examined in the Berlin Aging Study

	Parent		T1		T3		T4		T5		T6	
Variable	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
N	258	258	192	176	76	81	48	63	27	44	16	28
No. excluded	—	—	66	82	25	24	12	9	5	6	1	3
Age	84.73	85.11	83.64	82.54	84.23	83.66	83.26	84.28	85.27	86.05	88.55	89.57
	8.44	8.89	8.25	8.22	7.53	6.95	5.94	6.30	4.36	4.41	4.62	4.53
% aged 80+ years	67	67	63	56	55	61	63	71	96	93	100	100
Education, in years	11.31	10.19	11.69	10.43	12.01	10.81	11.92	10.87	12.11	11.18	11.88	10.71
	2.50	2.02	2.48	2.11	2.65	2.16	2.29	2.16	2.26	2.24	2.33	2.19
SMMSE	12.23	11.59	13.48	13.55	13.72	14.04	13.81	13.78	15.30	15.61	15.50	15.21
	2.61	3.37	1.25	1.33	1.22	1.01	1.07	1.14	1.94	1.85	1.86	2.13
Digit Letter	50.27	49.73	49.21	50.86	50.10	54.51	50.39	53.80	49.75	54.07	49.54	47.84
	9.14	10.80	9.46	10.52	9.64	8.93	10.42	8.85	10.70	10.21	9.38	9.69
Identical Pictures	50.89	49.11	50.22	49.77	53.58	53.77	52.82	54.69	55.38	57.83	53.21	50.59
	9.45	10.46	9.72	10.32	9.61	9.66	9.84	10.00	11.23	9.36	8.41	11.08
Paired Associates	49.37	50.63	48.51	51.64	49.70	53.50	50.25	54.32	50.67	56.67	48.63	52.32
	9.09	10.81	9.19	10.61	8.85	10.35	10.58	12.27	11.35	10.97	10.05	12.62
Memory for Text	49.24	50.76	48.45	51.69	48.98	52.85	51.06	52.30	52.00	55.71	51.85	52.16
	9.35	10.58	9.42	10.36	7.83	9.32	9.21	7.73	9.96	10.71	9.17	9.62
Categories	51.33	48.67	50.49	49.46	51.48	53.52	52.99	54.31	49.85	54.38	52.38	49.07
	9.23	10.56	9.29	10.72	10.25	10.23	10.30	12.45	10.73	12.70	12.10	13.65
Word Beginnings	50.32	49.68	49.40	50.66	50.85	54.25	52.72	54.80	52.11	54.39	50.04	50.94
	9.29	10.67	9.39	10.61	9.61	10.93	9.99	9.89	9.83	14.42	7.47	12.15
Vocabulary	51.72	48.28	51.55	48.32	51.79	49.84	56.00	53.99	57.36	55.99	57.83	51.12
	10.21	9.50	10.27	9.44	8.47	9.05	8.95	8.89	11.32	10.10	7.34	11.12
Spot-a-Word	50.28	49.72	49.44	50.59	51.76	51.65	50.34	52.15	46.68	53.24	52.13	52.69
	9.89	10.12	10.40	9.56	8.91	7.45	9.44	9.61	13.73	8.83	9.36	8.46

Notes: SMMSE = Short Mini-Mental State Exam (range = 0–18). For T1–T6, participants with suspected dementia were excluded from analyses. Means and standard deviations are shown. Cognitive tests represent *T* scores standardized to cross-sectional BASE parent sample ($N = 516$), $M = 50$, $SD = 10$.

known to be deceased from registry information. We assessed attrition effects on the cognitive tasks by using a method described by Lindenberger, Singer, and Baltes (2002). Participants who contributed two measurements performed on average 0.29 *SD* units higher at T1 than did participants assessed only once. Participants who subsequently contributed five occasions performed 0.53 *SD* higher at T1. These effects were higher for women than for men (e.g., T4: 0.47 *SD* vs 0.36 *SD*). On average, longitudinal participants had a higher education level (0.11 *SD*) and, as to be expected, they were younger at T1 (0.69 *SD*).

Measures

Cognition.—We used eight tests to assess four intellectual abilities: Digit Letter and Identical Pictures tests to assess perceptual speed; Paired Associates and Memory for Text tests to assess episodic memory; Categories and Word Beginnings tests to assess fluency; and Vocabulary and Spot a Word tests to assess knowledge. At each occasion, testing was supported by a Macintosh SE/30 computer equipped with a touch-sensitive screen; the same test versions were administered in individual sessions at the participant's residence (see Lindenberger & Baltes, 1997; Lövdén et al., 2004). Missing data because of poor vision amounted to 3% (205 of 6,008 data points).

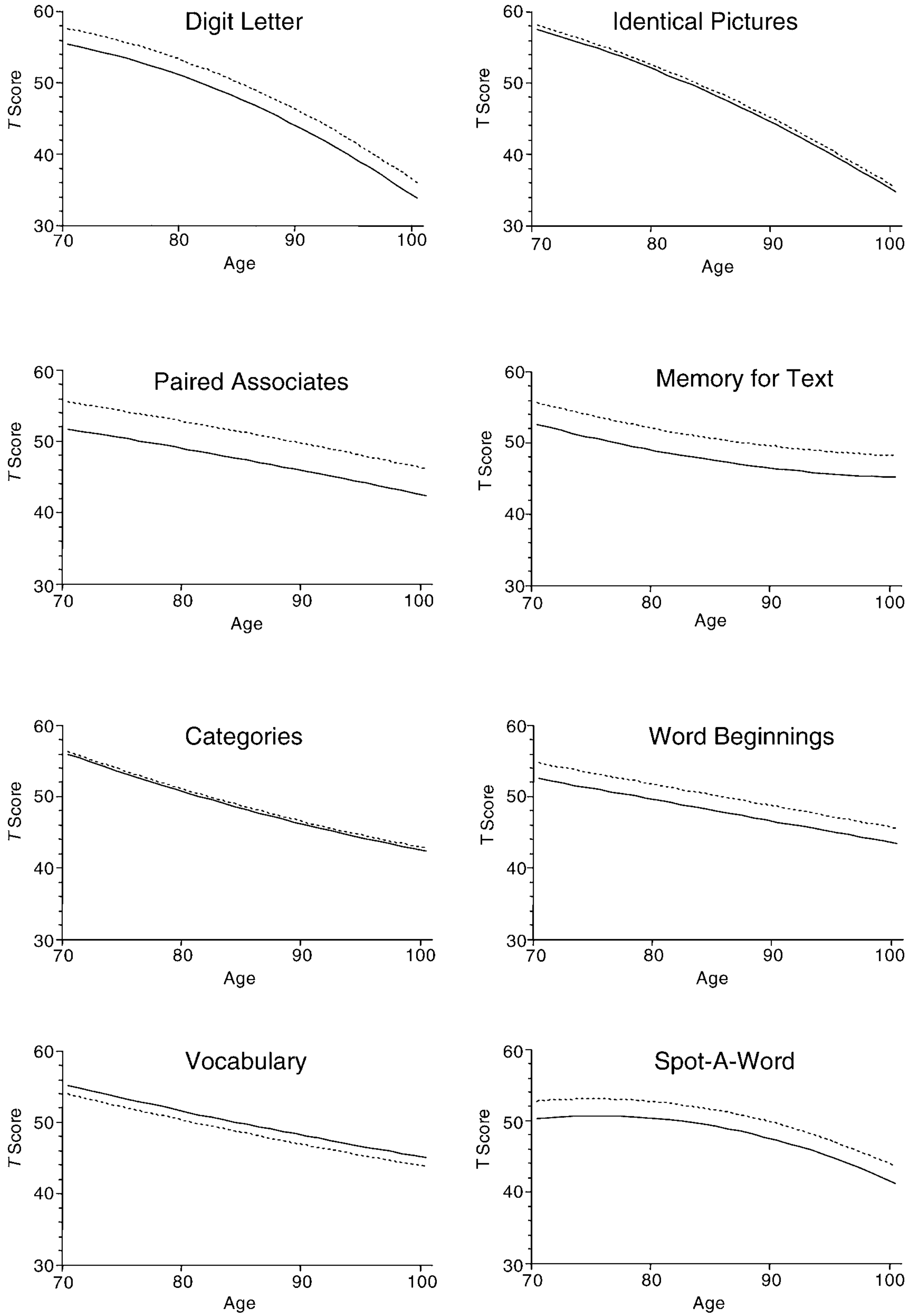
Education.—Educational information was not available for 12 participants, but we estimated it with linear regression analyses using equivalent income (see Baltes & Mayer, 1999). On average, men had more education (11.7 years, $SD = 2.5$, T scores $= 53.7$) than did women (10.4 years, $SD = 2.1$, T scores $= 48.5$): $F(1, 366) = 27.38$, $p < .001$.

RESULTS

We used the SAS procedure Proc Mixed (Littell, Miliken, Stoup, & Wolfinger, 1996) to estimate individual growth models from 70 to 100 years, separately for the eight cognitive tests. We centered age at 85 years for the estimation of intercepts, slopes, and intercept–slope covariances. We used the full-information maximum likelihood estimation algorithm, employed an unstructured covariance matrix for the random effects, and included all data available (5,803 observations, in total). This accommodated for unbalanced data structures, that is, individual differences in intervals between assessments, and incomplete data under the missing-at-random assumption (McArdle, 1994).

→

Figure 1. Intellectual ability age gradients residualized for education (using regression analyses), as observed in the Berlin Aging Study (BASE) as a function of sex. Solid lines represent men and dashed lines represent women. At the zero-order level, there were sex differences in intercepts of cognitive functioning favoring women on the Paired Associates and Memory for Text tests and those favoring men on the Vocabulary test. Once we took sex differences in education into account, the advantage for men vanished and women outperformed men on five out of the eight



cognitive tests: Digit Letter, Paired Associates, Memory for Text, Word Beginnings, and Spot a Word. For example, women outperformed men on the Paired Associates test by 4 *T*-score units. Although both show linear decline amounting to approximately 3 *T*-score units per decade, sex differences remain constant during advanced old age. We standardized *T* scores to the T1 parent BASE sample ($N = 516$), $M = 50$, $SD = 10$, to ensure a common metric while maintaining both the psychometric properties of the scores and the longitudinal changes in means and variances.

In a first step, we estimated the amount of between- and within-person variance by considering models that allowed random effects only for the intercept (i.e., no change allowed). Most of the total variation was between-person variance (61%, averaged across the tests), but each measure showed sizeable within-person variation that was indicative of change over time (e.g., Memory for Text test: 57%).

In a second step, we tested models including fixed effects for intercept, linear slopes, and quadratic slopes as well as random effects for intercept, linear slope, and the intercept–linear slope covariance (see endnote 2). Estimated fixed effects for all cognitive tests produced intercepts similar to the mean of the T1 sample (e.g., Digit Letter test: 49.93, $SE = 0.53$). Over time, all cognitive tests displayed negative linear age slopes (e.g., Digit Letter test: –0.78 *T*-score decline per year), and we found (marginally) significant negative quadratic age trends for six tests (the Digit Letter, Identical Pictures, Paired Associates, Word Beginnings, Vocabulary, and Spot-a-Word tests), suggesting an acceleration of linear decline among the oldest-old individuals. Descriptively, this component was strongest for the Digit Letter test (–0.03). We found random effects for the intercepts on all eight tests, whereas we found random effects for the linear slope only for the Digit Letter and Categories tests. The intercept–linear slope covariances were significant and positive for the Digit Letter, Identical Pictures, and Spot-a-Word tests.

To examine sources of variability, we entered education and gender as covariates. Education was significantly associated with variance in intercept on all tests (e.g., 0.37 for the Vocabulary test). Most importantly, once we took differences in education into account, women outperformed men on five out of eight tests: Digit Letter (2.19, $p < .05$), Paired Associates (4.09, $p < .001$), Memory for Text (3.54, $p < .001$), Word Beginnings (2.32, $p < .05$), and Spot-a-Word (2.22, $p < .05$). These effect sizes were between 2 and 4 *T*-score units. According to statistical convention (e.g., Cohen, 1988), the effects range from small (Spot a Word test) to almost moderate (Paired Associates test) in magnitude. None of the interactions (Education × Linear slope, Sex × Linear slope, and Education × Sex × Linear slope) was significantly different from zero, indicating preserved sex- and education-related differences during advanced old age (see endnote 3). Interestingly, analyses *not* adjusting for education indicated fewer advantages for women (Paired Associates test: 3.00, $p < .01$; Memory for Text test: 2.89; $p < .01$), as well as an advantage for men on the Vocabulary test (–2.89; $p < .01$). Figure 1 separately illustrates cognitive age trajectories for women and men after we residualized performance scores for education.

Discussion

Our goal was to extend longitudinal evidence on patterns of sex differences in cognitive aging during advanced old age. Using 13-year longitudinal data from 70- to 100-year-old participants in the BASE who were screened for dementia, we found that, after we controlled for gender- and cohort-related differences in education, women outperformed men on tests across the four cognitive domains examined. Over time, all cognitive tests showed linear (or even quadratic) decline over the 30-year period of old age modeled. Women and men showed parallel trajectories of cognitive aging, which suggests that neither sex nor education is associated with differential cognitive decline (see endnote 4).

As is well established in the cognitive aging literature, sample attrition was selective in that the more data participants provided for longitudinal change estimates, the better was their cognitive performance at T1. We acknowledge that the capability of likelihood analyses under the missing-at-random assumption to accommodate incomplete data (e.g., including attrition-informative variables such as age) is restricted if unobserved change considerably differs from observed change; we also acknowledge that our results are contingent on common, but untested, statistical assumptions (e.g., ergodicity; see Borsboom, Mellenbergh, & van Heerden, 2003). Indeed, our findings likely underestimate true cognitive change and associated interindividual differences. Future research should specifically explore whether sex-differential selection effects at the population level (e.g., survival, dementia) and at the sample level (e.g., attrition) nullify, create, or diminish differences between women and men (Smith & Baltes, 1998).

Life course theories suggest and evidence has shown that the negative consequences of fewer lifetime educational and occupational opportunities have followed current cohorts of old women into advanced old age. In line with this, our analyses suggest that men's overall higher level of education suppresses women's advantage on tasks assessing perceptual speed, episodic memory, fluency, and knowledge. Thus, independent of the biological underpinning, sociocultural factors appear to shape the magnitude of sex differences in cognitive functioning over the entire life span.

We found that, both when we did adjust and when we did not adjust for differences in early-life education, sex differences in cognition were stable into the highest ages with no evidence of change in magnitude. This sex-equivalent longitudinal decline is in line with recent evidence of parallel cortical brain atrophy in women and men (Raz et al., 2005). Although some research from major large-scale longitudinal studies have documented more rapid decline for women than for men (e.g., Zelinski & Stewart, 1998), several studies report sex differences in level, but not in subsequent, change (e.g., Finkel et al., 2003). Our findings add to this. It seems that the major biological and psychological changes associated with aging do not modify or affect sex differences in cognition. Given the extant interest in the heterogeneity of cognitive aging, a more promising approach for future studies may involve examining how macrostructural indices such as gender and education are involved in bringing about diverse within-domain and across-domain profiles of functioning in advanced old age (e.g., Smith & Baltes, 1998).

Acknowledgments

From 1988 until 1998, BASE was financially supported by two German federal departments: the Department of Research and Technology (13 TA 011: 1988–1991) and the Department of Family and Senior Citizens (1991–1998). Since 1999, BASE has been funded by the Max Planck Institute for Human Development, Berlin, where the study is located. Members of the Steering Committee are P. B. Baltes and J. Smith (psychology), K. U. Mayer (sociology), E. Steinhagen-Thiessen and M. Borchelt (internal medicine and geriatrics), and H. Helmchen and F. Reischies (psychiatry). Field research was coordinated at various phases by R. Nuthmann, M. Neher, and K. Fröhlich.

Denis Gerstorf is now in the Psychology Department at the University of Virginia, on a Research Fellowship awarded by the German Research Foundation (DFG). This article was prepared while Agneta Herlitz was visiting scholar to the Max Planck Institute for Human Development, Berlin, Germany.

Address correspondence to Denis Gerstorf, Psychology Department, University of Virginia, P.O. Box 400400, Charlottesville, VA 22904-4400. E-mail: gerstorf@virginia.edu

References

Baltes, P. B., & Mayer, K. U. (Eds.). (1999). *The Berlin Aging Study: Aging from 70 to 100*. New York: Cambridge University Press.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review, 110*, 203–219.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Erlbaum.

Elias, M. F., Elias, P. K., D'Agostino, R. B., Silbershatz, H., & Wolf, P. A. (1997). Role of age, education, and gender on cognitive performance in the Framingham Heart Study: Community-based norms. *Experimental Aging Research, 23*, 201–235.

Finkel, D., Reynolds, C. A., McArdle, J. J., Gatz, M., & Pedersen, N. L. (2003). Latent growth curve analyses of accelerating decline in cognitive abilities in late adulthood. *Developmental Psychology, 39*, 535–550.

Fratiglioni, L., Viitanen, M., von Strauss, E., Tontodonati, V., Herlitz, A., & Winblad, B. (1997). Very old women are at the highest risk for dementia and Alzheimer's disease: Incidence data from the Kungsholmen Project, Stockholm. *Neurology, 48*, 132–138.

Hassing, L., Wahlin, A., & Bäckman, L. (1998). Minimal influence of age, education, and gender on episodic memory functioning in very old age: A population-based study of nonagenarians. *Archives of Gerontology and Geriatrics, 27*, 75–87.

Klein, L. E., Roca, R. P., McArthur, J., Vogelsang, G., Klein, G. B., Kirby, S. M., et al. (1985). Diagnosing dementia. Univariate and multivariate analyses of the Mental Status Examination. *Journal of the American Geriatrics Society, 33*, 483–488.

Lindenberger, U., & Baltes, P. B. (1997). Intellectual functioning in old and very old age: Cross-sectional results from the Berlin Aging Study. *Psychology and Aging, 12*, 410–432.

Lindenberger, U., Singer, T., & Baltes, P. B. (2002). Longitudinal selectivity in aging populations: Separating mortality-associated versus experimental components in the Berlin Aging Study (BASE). *Journals of Gerontology: Psychological Sciences, 57B*, P474–P482.

Littell, R. C., Miliken, G. A., Stoup, W. W., & Wolfinger, R. D. (1996). *SAS system for mixed models*. Cary, NC: SAS Institute.

Lövdén, M., Ghisletta, P., & Lindenberger, U. (2004). Cognition in the Berlin Aging Study (BASE): The first 10 years. *Aging, Neuropsychology, and Cognition, 11*, 104–133.

Maitland, S. B., Herlitz, A., Nyberg, L., Bäckman, L., & Nilsson, L.-G. (2004). Selective sex differences in declarative memory. *Memory and Cognition, 32*, 1160–1169.

McArdle, J. J. (1994). Structural factor analysis experiments with incomplete data. *Multivariate Behavioral Research, 29*, 409–454.

Meinz, E. J., & Salthouse, T. A. (1998). Is age kinder to females than to males? *Psychonomic Bulletin and Review, 5*, 56–70.

Rabbitt, P., Lunn, M., & Wong, D. (2005). Neglect of dropout underestimates effects of death in longitudinal studies. *Journals of Gerontology: Psychological Sciences, 60B*, P106–P109.

Raz, N., Lindenberger, U., Rodriguez, K. M., Kennedy, K. M., Head, D., Williamson, A., et al. (2005). Regional brain changes in aging healthy adults: General trends, individual differences, and modifiers. *Cerebral Cortex, 15*, 1676–1689.

Schaie, K. W. (2005). Developmental influences on adult intelligence: The Seattle Longitudinal Study. Oxford, England: Oxford University Press.

Seeman, T. E., Huang, M.-H., Bretsky, P., Crimmins, E., Launer, L., & Guralnik, J. M. (2005). Education and APOE-e4 in longitudinal cognitive decline: MacArthur Studies of Successful Aging. *Journals of Gerontology: Psychological Sciences, 60B*, P74–P83.

Singer, T., Verhaeghen, P., Ghisletta, P., Lindenberger, U., & Baltes, P. B. (2003). The fate of cognition in very old age: Six-year longitudinal findings in the Berlin Aging Study (BASE). *Psychology and Aging, 18*, 318–331.

Sliwinski, M., & Buschke, H. (2004). Modeling intra-individual cognitive change in aging adults: Results from the Einstein Aging Studies. *Aging, Neuropsychology, and Cognition, 11*, 196–211.

Smith, J., & Baltes, M. M. (1998). The role of gender in very old age: Profiles of functioning and everyday life patterns. *Psychology and Aging, 13*, 676–695.

Zelinski, E., & Stewart, S. T. (1998). Individual differences in 16-year memory changes. *Psychology and Aging, 13*, 622–630.

Received July 19, 2005
Accepted December 3, 2005
Decision Editor: Thomas M. Hess, PhD

End Notes

1. Although we did not exclude significantly more women than men, $\chi^2(1, N = 516) = 2.43$, $p > .10$, sex differences among those excluded were found in age—men, 87.9 years; women, 90.6 years; $F(1, 146) = 4.19$, $p < .05$—and SMMSE—men, $M = 8.6$; women, $M = 7.4$; $F(1, 146) = 9.77$, $p < .01$. Women and men included in the study did not differ in age, T1, $F(1, 366) = 1.62$, $p > .10$, and SMMSE, T1, $F(1, 366) = 0.21$, $p > .10$.
2. We residualized the quadratic terms for linear terms of age (as obtained in regression analyses; cf. Lövdén et al., 2004) to arrive at a quadratic component that is independent of the linear component. Follow-up analyses of these quadratic terms revealed that none of the eight cognitive tests displayed statistically significant individual differences in the quadratic age trends, thereby precluding the possibility to test for sex differences in quadratic age trends.
3. In follow-up analyses, we additionally included flat or linear retest effects in our models. Analogous to the findings of Lövdén and colleagues (2004), we found evidence for either or both types of retest effects on some of the tests (e.g., Memory for Text). However, the reported pattern of cognitive sex differences remained virtually the same. We also modeled cognitive change over time in study (see also Sliwinski & Buschke, 2004) and used chronological age, education, and sex as covariates. Results of these analyses again were consistent with the pattern reported.
4. Although our data basis would have, in principle, made it possible to consider higher order relations above quadratic functions, attrition substantially reduced our sample of participants contributing 3 or more data points, thereby making it difficult to reliably test for other than linear and quadratic age functions. Furthermore, we found individual differences only in the linear component of age-related change, which precluded us from being able to test for sex differences in other than the linear component.