

Influence of Age on Practice Effects in Longitudinal Neurocognitive Change

Timothy A. Salthouse
University of Virginia

Objective: Longitudinal comparisons of neurocognitive functioning often reveal stability or age-related increases in performance among adults under about 60 years of age. Because nearly monotonic declines with increasing age are typically evident in cross-sectional comparisons, there is a discrepancy in the inferred age trends based on the two types of comparisons. The current research investigated the role of practice effects in longitudinal comparisons on the discrepancy. **Method:** Longitudinal data over an average interval of 2.5 years were available on five abilities (i.e., reasoning, spatial visualization, episodic memory, perceptual speed, vocabulary) in a sample of 1,616 adults ranging from 18 to over 80 years of age. Practice effects were estimated from comparisons of the performance of people of the same age tested for either the first or second time, after adjusting for the possibility of selective attrition. **Results:** Increased age was associated with significantly more negative longitudinal changes with each ability. All of the estimated practice effects were positive, but they varied in magnitude across neurocognitive abilities and as a function of age. After adjusting for practice effects the longitudinal changes were less positive at younger ages and slightly less negative at older ages. **Conclusions:** It was concluded that some, but not all, of the discrepancy between cross-sectional and longitudinal age trends in neurocognitive functioning is attributable to practice effects positively biasing the longitudinal trends. These results suggest that the neurobiological substrates of neurocognitive functioning may change across different periods in adulthood.

Keywords: retest effects, cross-sectional versus longitudinal, within-individual neurocognitive change

There is often a striking discrepancy in the age trends obtained from cross-sectional and longitudinal comparisons of certain cognitive abilities, as the former frequently reveal nearly linear declines starting in the early twenties, whereas significant longitudinal decline seldom occurs until adults are 60 years of age or older. Many factors have been speculated to contribute to this discrepancy (e.g., Baltes, Reese, & Nesselroade, 1977; Salthouse, 2010). Among the most frequently mentioned are cohort effects associated with time-related changes in culture or society that affect cross-sectional trends, and practice or retest effects that affect longitudinal trends.

Although none by itself is definitive, several characteristics of the discrepancy seem consistent with the possibility that at least a portion of it is attributable to practice effects distorting longitudinal trends. For example, the discrepancy is present in nonhuman animals raised in controlled environments for whom cohort influences are unlikely (e.g., Algeri, Biagini, Manfredi, & Pitsikas, 1991; Caprioli, Ghirardi, Giuliani, Ramacci, & Angelucci, 1991; Dellu, Mayo, Vallee, Le Moal, & Simon, 1997; Markowska & Savonenko, 2002). However, the discrepancy is either absent, or

much reduced, in measures of perceptual speed (e.g., Schaie, 1989), and in measures of brain volume (e.g., Du et al., 2006; Fotenos, Snyder, Gorton, Morris, & Buckner, 2005; Liu et al., 2003; Raz et al., 2005; Scahill et al., 2003; Taki, Kinomura, Sato, Goto, Kawashima, & Fukuda, 2009), in which little or no practice effects associated with prior assessment might be expected.

Another relevant characteristic is that the discrepancy is more pronounced at younger ages because longitudinal change is often positive below about 60 years of age, but negative at older ages (see Figure 2.2 in Salthouse, 2010). This property is illustrated in Figure 1, which portrays longitudinal changes (i.e., T_2 score – T_1 score) as a function of age in two separate projects. The data in the left panel are changes over a 7-year interval (from Table 5.1 in Schaie, 2005), and the data in the right panel are changes over a 5-year interval (from Table 3 in Ronnlund et al., 2005, and Table 3 in Ronnlund & Nilsson, 2006). In each case, there is a nearly linear relation between age and the magnitude of longitudinal change, with positive change at younger ages and negative change at older ages. Because cross-sectional age differences are generally negative at all ages, the shifts in the direction of the longitudinal changes tend to result in a smaller discrepancy between cross-sectional and longitudinal age trends at older ages.

The existence of a systematic relation between age and longitudinal change raises the question of the relation between age and the magnitude of practice effects in longitudinal comparisons. Specifically, are the benefits of prior experience similar across adulthood, such that the observed longitudinal changes are positively biased at all ages, or are the positive changes apparent before about 50 or 60 years of age attributable to greater effects of prior testing experience at younger ages? Because the answers to these

Timothy A. Salthouse, Department of Psychology, University of Virginia.

I acknowledge the contributions of the research assistants and research participants in the Virginia Cognitive Aging Project (VCAP), which is the source of data in this report. The research was supported by NIA Grant R37AG024270.

Correspondence concerning this article should be addressed to Timothy A. Salthouse, Department of Psychology, 102 Gilmer Hall, University of Virginia, Charlottesville, VA 22904-4400. E-mail: salthouse@virginia.edu

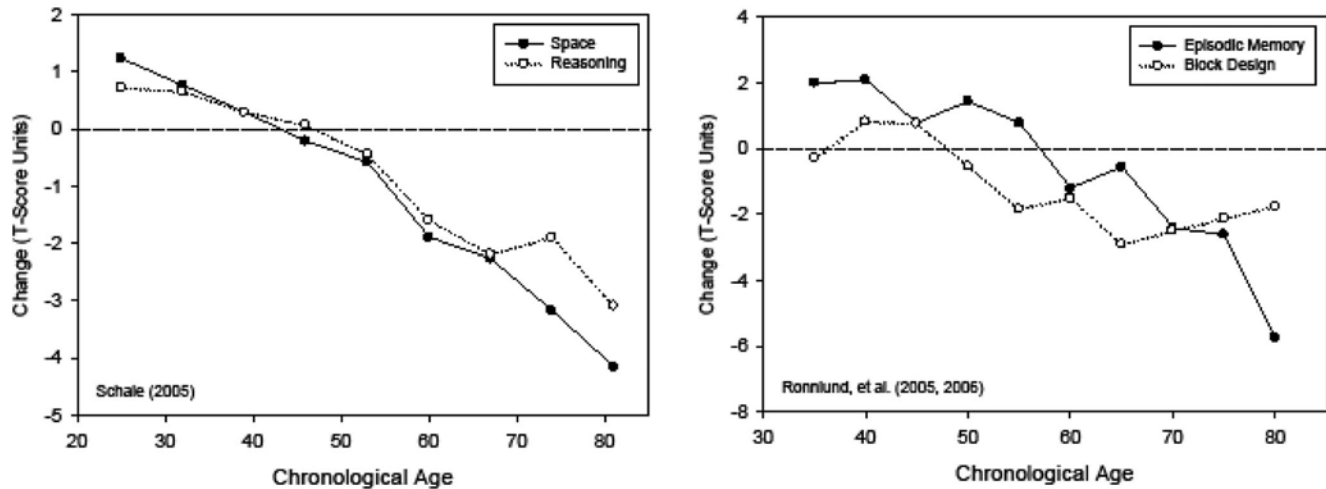


Figure 1. Mean longitudinal change ($T_2 - T_1$) in two variables from Schaie (2005) and two variables from Ronnlund and colleagues (2006, 2005).

questions are relevant to whether age-related neurocognitive change occurs continuously across all of adulthood or abruptly after a particular age, they are fundamental to understanding the nature of cognitive aging. The role of practice on longitudinal change is also relevant to understanding the neurobiological substrates of cognitive change. That is, if practice effects dominate at younger ages then there may be qualitative shifts in the relative contributions of different neurobiological substrates of cognitive change at different ages. In contrast, age-invariant influences of the neurobiological substrates might be expected if practice effects were found to be nearly constant at all ages.

The primary goal of the current project was to investigate the influence of practice effects on longitudinal changes in neurocognitive functioning, and to determine if the effects are greater at younger ages. Positive practice effects have been reported among young adults (e.g., Ronnlund & Nilsson, 2006; Ronnlund et al., 2005; Salthouse, 2009a; Salthouse, Schroeder, & Ferrer, 2004), but the results of previous studies have been mixed regarding the relations of age to the magnitude of practice effects (e.g., Ferrer, Salthouse, McArdle, Stewart, & Schwartz, 2005; Ferrer, Salthouse, Stewart, & Schwartz, 2004; Wilson et al., 2002; Wilson, Li, Bienias, & Bennett, 2006). The inconsistencies across studies may be attributable to variations in the cognitive abilities that were examined, and in the age range of the samples of research participants. Some of these inconsistencies may be resolvable in the current project because it involved the examination of five cognitive abilities in adults across a wide age range.

At least three different methods have been used to derive estimates of practice effects in longitudinal research. One method can be used when three or more assessments are available from each participant such that influences associated with age and influences associated with test experience can be modeled with different functions (e.g., Ferrer et al., 2005, 2004; Finkel, Reynolds, McArdle, & Pedersen, 2005; Ghisletta & de Ribaupierre, 2005; Lovden, Ghisletta, & Lindenberger, 2004; Rabbitt, Diggle, Holland, & McInnes, 2004; Rabbitt, Diggle, Smith, Holland, & McInnes, 2001; Rodgers, Ofstedal, & Herzog, 2003; Wilson et al.,

2002, 2006). A second method is based on a design with retest intervals which vary across participants, such that there is no longer a perfect correlation between the increase in age and the increase in test experience, thereby allowing separate estimates of the two influences to be derived (e.g., McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002; Salthouse, 2009a; Salthouse et al., 2004).

A third method of estimating practice effects in longitudinal research is based on a comparison of the performance of individuals tested for the second time with the performance of people of the same age who are tested only once. Although seemingly straightforward, the comparison at the second occasion may be biased if the two samples were not equivalent at the initial occasion. That is, individuals who return for a second assessment are often positively selected from the initial sample (e.g., Brayne et al., 1999; Christensen et al., 2004; Deary & Der, 2005; Hultsch, Hertzog, Dixon, & Small, 1998; Kennison & Zelinski, 2005; Owens, 1953; Prins et al., 2005; Schaie, Labouvie, & Barrett, 1973; Singer, Verhaeghen, Ghisletta, Lindenberger, & Baltes, 2003), and thus people with and without prior test experience may also differ in their degree of selectivity relative to the people in the total sample who were only tested once.

One possible solution to the selective attrition problem is to specify the complete sample of all potential participants at the first occasion, and obtain some initial information from everyone at that time. The cognitive assessment of primary interest could then be administered to a randomly selected half of the participants at the first, T_1 , occasion and to everyone at the second, T_2 , occasion. Although a design such as this should allow effects of selective attrition to be assessed with respect to the initial information, and if necessary corrected for in both groups, there are apparently no published reports in which it has been implemented.

An alternative approach to the selective attrition problem consists of evaluating the degree of selectivity of the returning (longitudinal) sample relative to the initial sample, and adjusting the scores of individuals with no prior testing experience to account for selective attrition. Ronnlund et al. (2005; also see Schaie, 2005)

described a procedure of this type in which the difference (D) between scores of people tested for a second time and those tested for a first time was assumed to be composed of effects of test practice (P) and of selective attrition (A). The D values in their project were estimated by a contrast between scores of people of the same age and tested in the same year for either the first or second time, and the value of A was estimated from the difference in scores at the first occasion between the entire sample and the longitudinal sample who returned for a second assessment. Subtraction of A from D therefore yields an estimate of the practice effect adjusted for any initial differences in selectivity.

The current project used a version of the preceding method to derive estimates of practice effects, and of practice-adjusted longitudinal changes, in five cognitive abilities in adults across a wide range of ages. The data were obtained from a longitudinal study involving adults who were between 18 and 97 years of age at the initial occasion, with continuous recruitment of new participants from 2001 to 2009. Longitudinal retests began in 2004, with the interval between tests deliberately varied across participants, resulting in a range from 1 to 8 years and an average interval of 2.5 years. The initial, T1, assessments for the 3,782 individuals tested at least once occurred between 2001 and 2009, and the second, T2, assessments for the 1,616 individuals tested at least twice were between 2004 and 2009. Some of the participants were tested three or more times, but only the first and second assessments are considered in the current report.

Sixteen variables were selected to represent five cognitive abilities. Because the variables had a similar factor structure at different ages (Salthouse, 2004), the z-scores for the three or four tests representing each factor were averaged to form composites for each ability.

Four means were derived in each of 7 decade groups for each of the five composite scores: M1 = the mean of the total sample at T1; M2 = the mean of the subsample of participants only tested once at T1; M3 = the mean of the longitudinal sample at T1; and M4 = the mean of the longitudinal sample at T2. Following the procedures of Ronnlund et al. (2005) and Schaie (2005), these four means were used to derive the following estimates for each ability in each age group: D = M4 – M2 (difference between the second test for the longitudinal sample and the first test for the sample tested only once); A = M3 – M1 (selectivity of the longitudinal sample relative to the total sample); P = D – A (practice effect adjusted for selectivity); Change = M4 – M3 (longitudinal change); and Adjusted Change = Change – P (longitudinal change adjusted for the practice estimate).

Recruitment and testing in the Ronnlund et al. (2005) and Schaie (2005) projects occurred in discrete years, and thus the new and returning samples could be compared in the same year. Because recruitment and testing in the current project was continuous, the same types of comparisons in specific years were not feasible. Instead, the comparisons were based on data collapsed across all testing years, and potential period effects associated with the year

Table 1
Demographic Characteristics of the Total Sample and the Longitudinal Subsample by Decade

Characteristics	All	18–29	30–39	40–49	50–59	60–69	70–79	80–89
<i>N</i>								
Total	3,782	688	351	602	847	600	461	233
Longitudinal	1,616	188	136	282	393	296	234	87
<i>Age</i>								
Total	51.1 (18.6)	23.0 (3.1)	34.3 (2.8)	45.0 (2.9)	54.4 (2.8)	64.2 (2.9)	74.2 (2.8)	84.0 (3.7)
Longitudinal	53.7 (17.2)	22.2 (3.3)	34.8 (2.9)	45.2 (2.9)	54.5 (2.9)	64.4 (2.8)	74.1 (2.8)	83.7 (3.6)
<i>Proportion female</i>								
Total	.65	.57	.72	.73	.70	.67	.59	.48
Longitudinal	.65	.56	.70	.73	.71	.64	.59	.46
<i>Health rating</i>								
Total	2.2 (0.9)	2.0 (0.9)	2.1 (0.8)	2.1 (0.9)	2.2 (0.9)	2.1 (0.9)	2.4 (0.9)	2.6 (0.8)
Longitudinal	2.2 (0.9)	2.0 (0.9)	2.3 (0.8)	2.1 (0.9)	2.1 (0.9)	2.1 (0.8)	2.4 (0.9)	2.6 (0.9)
<i>Years of education</i>								
Total	15.6 (2.7)	14.7 (2.1)	15.8 (2.8)	15.3 (2.6)	15.9 (2.6)	16.4 (2.8)	15.9 (2.8)	16.2 (3.2)
Longitudinal	15.7 (2.7)	14.0 (1.8)	15.5 (2.5)	15.4 (2.4)	16.0 (2.6)	16.4 (2.6)	15.9 (2.9)	16.4 (3.4)
<i>MMSE</i>								
Total	28.5 (1.8)	29.0 (1.4)	28.7 (1.6)	28.6 (1.7)	28.6 (1.7)	28.6 (1.7)	28.3 (1.7)	27.1 (2.5)
Longitudinal	28.6 (1.6)	28.7 (1.6)	28.4 (1.7)	28.5 (1.7)	28.9 (1.4)	28.8 (1.4)	28.4 (1.7)	27.5 (2.1)
<i>Vocabulary Scaled score</i>								
Total	12.6 (3.0)	13.2 (3.1)	11.8 (3.4)	11.5 (3.1)	12.3 (3.0)	13.1 (2.4)	12.9 (3.0)	13.4 (2.9)
Longitudinal	12.8 (3.0)	12.8 (3.3)	11.6 (3.1)	11.8 (2.9)	12.8 (2.9)	13.5 (2.2)	13.3 (3.1)	14.1 (2.7)
<i>Digit Symbol Scaled score</i>								
Total	11.3 (2.9)	11.2 (2.8)	11.5 (3.1)	10.7 (2.9)	11.5 (2.9)	11.4 (2.7)	11.7 (2.9)	11.5 (2.8)
Longitudinal	11.5 (2.8)	10.8 (2.6)	11.4 (3.1)	11.1 (3.0)	11.7 (2.9)	11.6 (2.4)	12.0 (2.8)	12.0 (2.6)
<i>Word Recall Scaled score</i>								
Total	12.2 (3.3)	12.3 (3.1)	11.9 (3.2)	12.3 (3.4)	12.6 (3.2)	12.6 (3.2)	12.2 (3.3)	11.1 (3.6)
Longitudinal	12.5 (3.3)	11.8 (3.1)	11.7 (3.5)	12.2 (3.6)	12.8 (3.1)	13.2 (3.1)	12.4 (3.2)	11.8 (3.3)
<i>Logical Memory Scaled score</i>								
Total	11.8 (2.9)	11.7 (2.8)	11.4 (2.8)	11.2 (2.9)	11.8 (2.9)	12.3 (2.9)	12.2 (2.8)	11.9 (3.0)
Longitudinal	12.0 (2.8)	11.2 (2.8)	11.5 (3.0)	11.4 (2.7)	12.2 (2.8)	12.7 (2.7)	12.5 (2.8)	12.6 (2.7)
<i>Retest interval (years)</i>								
Total	2.5 (1.1)	2.4 (0.9)	2.5 (1.0)	2.8 (1.3)	2.6 (1.0)	2.5 (1.0)	2.6 (1.1)	2.3 (0.9)

Note. MMSE = Mini-Mental State Examination. Health was rated on a scale ranging from 1 = "excellent" to 5 = "poor."

of first testing were investigated with regression equations. All of the effects of the T1 year on the cognitive ability composite scores were small, but linear relations of year of first test were nevertheless removed from all scores prior to subsequent analyses.

Some of the participants in the project received three versions of the same tests across three sessions in the initial occasion, whereas other participants performed different tests across the three sessions in the initial occasion (cf., Salthouse, 2007). In order to maximize the sample size, only the score at the first session on each occasion was considered in the reported analyses. However, the influence of performing one or three versions at the initial occasion on scores at the second occasion was investigated by comparing the T2 scores for participants receiving three versus one version at the initial occasion after adjusting for T1 performance,

age, and retest interval. None of the standardized beta coefficients associated with number of test versions at T1 were significantly different from zero, and all were less than .01, and thus the number of versions at the initial occasion was ignored in subsequent analyses.

Method

Participants

All participants signed an informed consent form and the project was approved by the local institutional review board responsible for protection of human subjects. Table 1 contains demographic characteristics for the total sample and for the longitudinal subsample as a function of age decade. It can be seen that on average

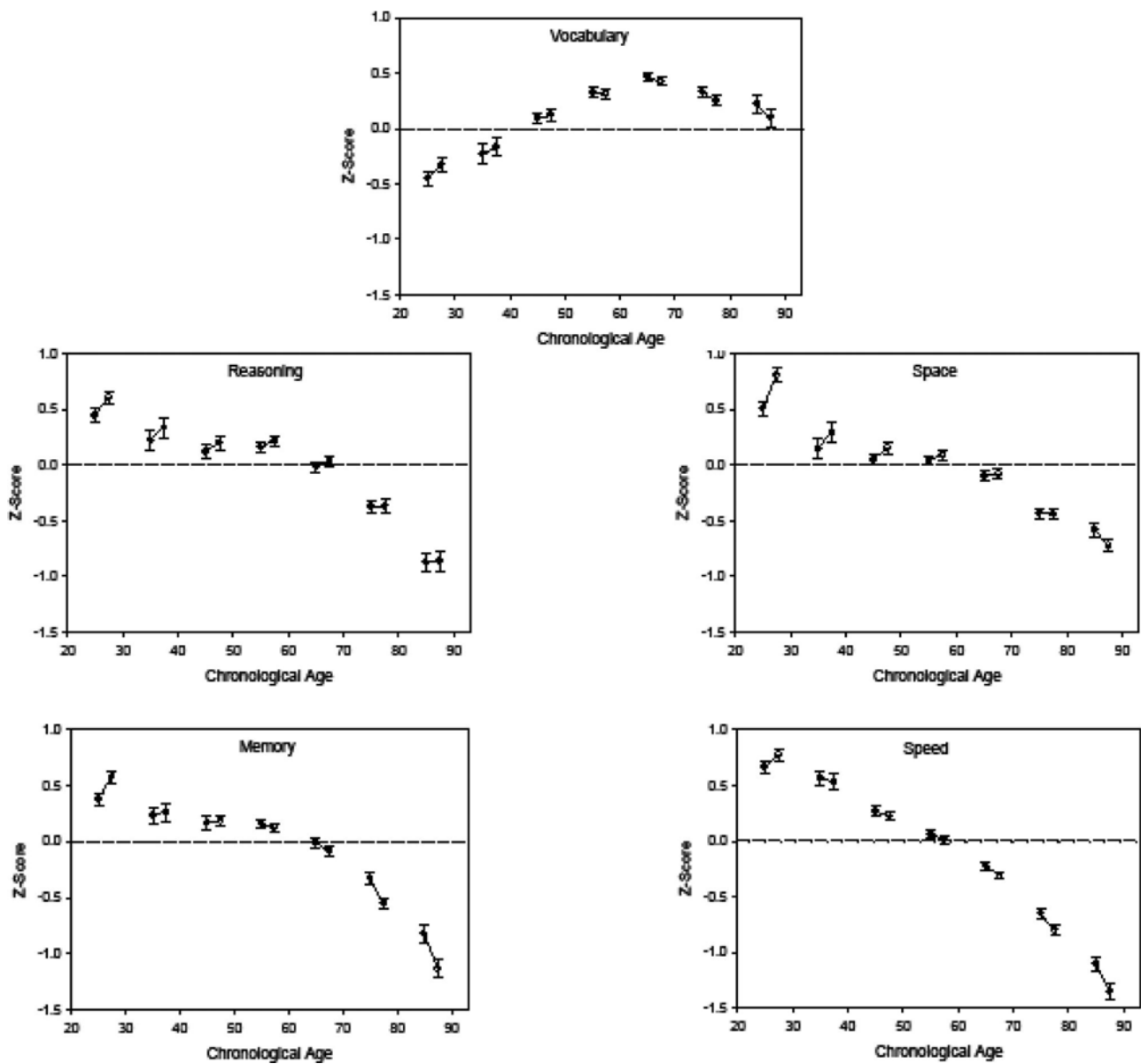


Figure 2. Mean performance on the first (T1) and second (T2) occasion by decade in five composite scores.

the participants rated their health in the good to very good range, and had completed about 15 years of education. The age-adjusted scaled scores are informative about the level of performance relative to the nationally representative samples used to establish the norms for the Wechsler Adult Intelligence Scale—Third Edition (WAIS III; Wechsler, 1997a) and the Wechsler Memory Scale—Third Edition (WMS III; Wechsler, 1997b). The means in the normative samples were set to 10 with a *SD* of 3, and therefore the means in the current samples were between .5 and 1 *SD* above the levels in the normative samples. However, two important points should be noted about these results. First, the standard deviations in each age decade were nearly the same magnitude as in the normative sample, indicating roughly comparable variability in these samples and in the nationally representative normative samples. And second, the age-adjusted values were similar at all ages, indicating nearly equivalent positive selection at all ages. It is also noteworthy that the age trends in the factor scores have been found to closely resemble those from nationally representative samples used to create norms for different commercial test batteries (i.e., Salthouse, 2009b, 2010).

Cognitive Tests

The cognitive tests are described in the Appendix, along with three estimates of their reliabilities. It is apparent that all of the variables had moderately high reliability as evaluated by internal consistency (i.e., coefficient alpha), alternate forms (i.e., correlation with other test versions), and test-retest (i.e., longitudinal stability coefficient) procedures. More details about the tests and administration procedures, as well as results of confirmatory factor analyses indicating the pattern of relations of variables to ability constructs, have been reported in other publications (Salthouse, 2004, 2005, 2007; Salthouse, Pink, & Tucker-Drob, 2008; Salthouse & Tucker-Drob, 2008). The three tests representing each cognitive ability were Matrix Reasoning, Shipley Abstraction, and Letter Sets for reasoning, Spatial Relations, Paper Folding, and Form Boards for spatial visualization, Word Recall, Paired Associates, and Logical Memory for memory, and Digit Symbol, Pattern Comparison, and Letter Comparison for perceptual speed. Four tests were used to assess vocabulary: Vocabulary, Picture Vocabulary, Synonym Vocabulary, and Antonym Vocabulary.

Results

The initial step in the analyses consisted of converting the scores of each test to *z* scores based on the distribution of scores in the entire sample on the first (T1) occasion, and then averaging them to form composite scores for each cognitive ability. The five panels of Figure 2 illustrate the longitudinal data for each composite score. The T1 scores are represented by solid circles, and the T2 scores for the same individuals after an average of 2.5 years are represented by open circles. It can be seen that there is a shift in the direction of longitudinal change, represented by the orientation of the line connecting the pair of circles, from positive to negative with increased age. The composite speed variable is an exception because the longitudinal and cross-sectional trends were similar at all but the youngest age.

Selectivity estimates corresponding to the mean of the longitudinal sample at T1 minus the mean of the total sample at T1 (i.e.,

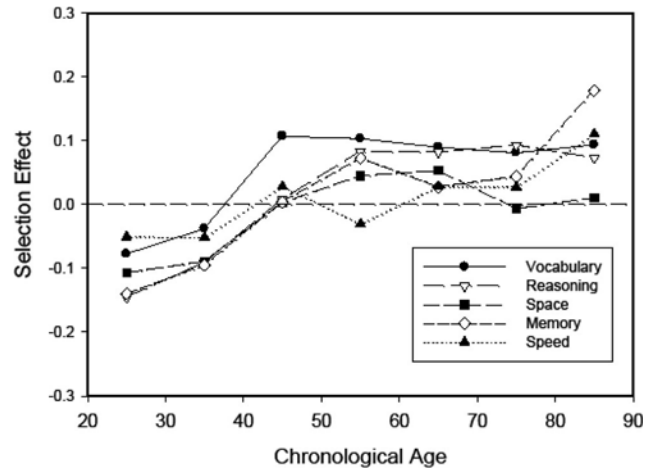


Figure 3. Estimates of selection effects in five composite scores as a function of age computed by subtracting the mean score of the complete sample from the mean score of the longitudinal sample on the first occasion.

values of *A*) are portrayed in Figure 3. Two features of these results are particularly noteworthy. First, the patterns were similar for all abilities, albeit with somewhat larger positive selectivity in memory at the oldest age. And second, the negative selectivity at the two youngest ages indicates that the young longitudinal participants had lower initial levels of functioning than the total sample of young participants. The negative selectivity in the young adults likely occurred because there was greater mobility of high functioning young professionals and students in the initial sample, and thus they were not available to participate in a second occasion. As is often found in longitudinal research with older adults, however, the older returners had slightly higher levels of functioning than the older sample as a whole.

The *D*, *A*, and *P* values were computed for each ability in each age decade according to the procedures described above. Estimates of longitudinal change ($T2 - T1$), practice ($D - A$), and adjusted change ($\text{Change} - P$) for each cognitive ability by decade are presented in the five panels of Figure 4. The age trends with the observed changes in Figure 4 are similar to those in Figure 1, but it should be noted that these changes are in *z* score units across a 2.5-year interval, whereas Figure 1 portrays changes in *T* score units (which have a standard deviation of 10 instead of 1 as in *z* scores) across either a 7-year (Schaie) or a 5-year (Ronnlund) interval.

The observed changes in every composite ability were negatively correlated with age (i.e., $-.21$ for vocabulary, $-.11$ for reasoning, $-.21$ for space, $-.25$ for memory, and $-.17$ for speed). Possible nonlinear trends in the changes were also examined, but all quadratic trends were associated with less than .002 proportion of the variance, and therefore there was no evidence of a discrete transition from stability to decline with increasing age.

It is apparent in Figure 4 that the estimated practice effects varied across cognitive abilities both in their absolute magnitude, and with respect to their relations with age. However, for most of the abilities the practice estimates were as large as, or larger than, the observed longitudinal changes. There was also variation in the

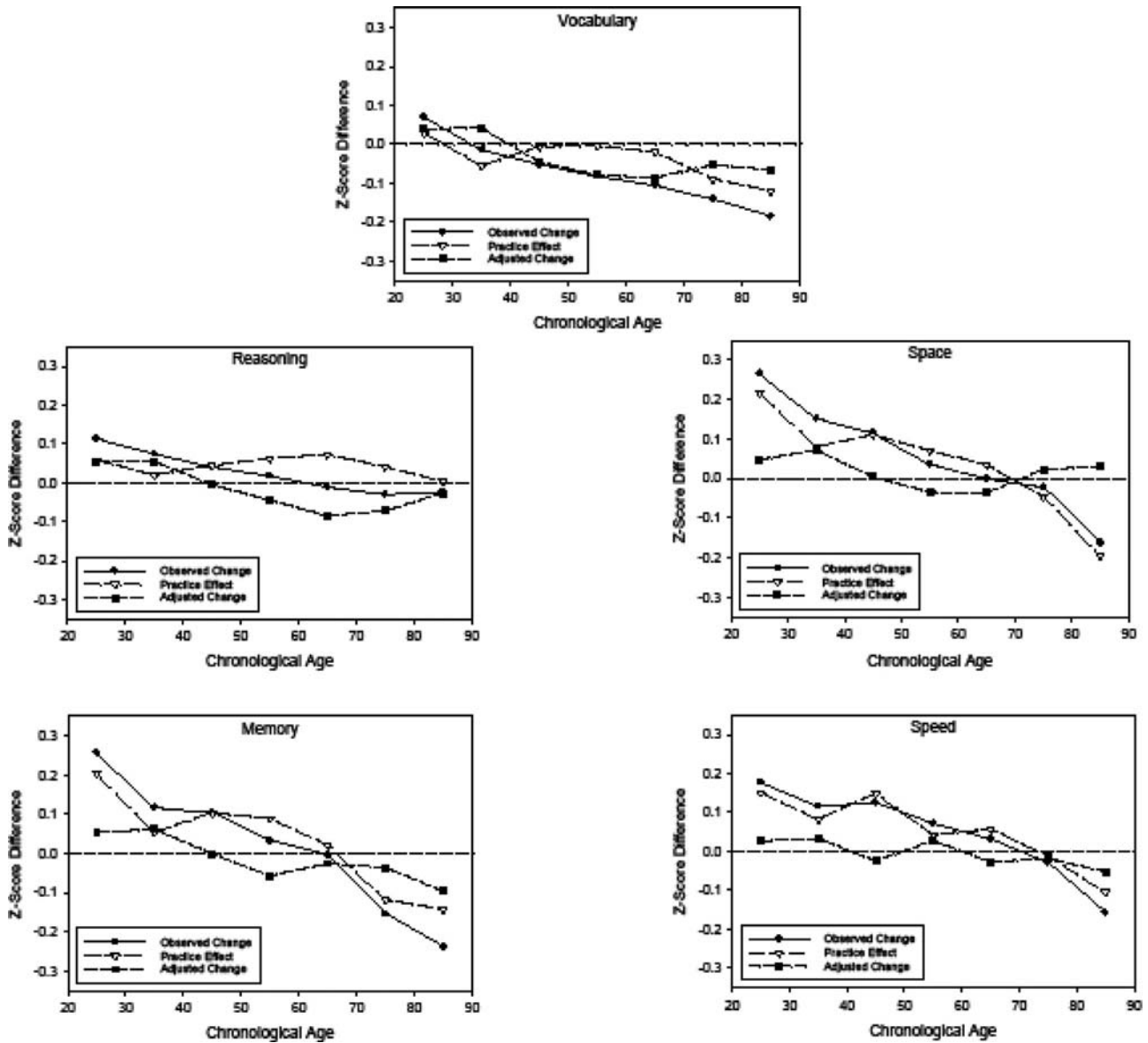


Figure 4. Observed longitudinal change and estimates of practice and practice-adjusted longitudinal change in five composite scores as a function of age.

age-practice relations as there was no relation of age on the practice estimates in reasoning ability, and the practice effects in vocabulary were very small and primarily restricted to the two oldest groups.

Subtraction of the estimated practice effects from the observed longitudinal changes resulted in a flattening of the age trends, with much less positive change at younger ages and slightly less negative change at older ages. An implication of these results is that longitudinal changes would likely be smaller, and less extreme in both positive and negative directions, if practice effects were not contributing to the changes.

Another method of portraying the data involves a direct contrast of cross-sectional differences with the observed, and the practice-adjusted, longitudinal changes. In order to provide stable esti-

mates, the sample was divided into only two groups for this analysis based on the median age in the longitudinal sample. The young group ($N = 765$) ranged from 18 to 53 years of age with a mean of 37, and the old group ($N = 851$) ranged from 54 to 97 years of age with a mean of 67.

Expected cross-sectional differences over a period comparable to the average longitudinal interval of 2.5 years were derived by multiplying the slope of the regression equation relating T1 score to age in the total sample of individuals within the appropriate age range by 2.5. These values, along with the observed longitudinal change, and the practice-adjusted longitudinal change, are plotted in Figure 5 for each ability in the two age groups.

Inspection of Figure 5 reveals that for the young adults there was a large discrepancy between the cross-sectional age difference

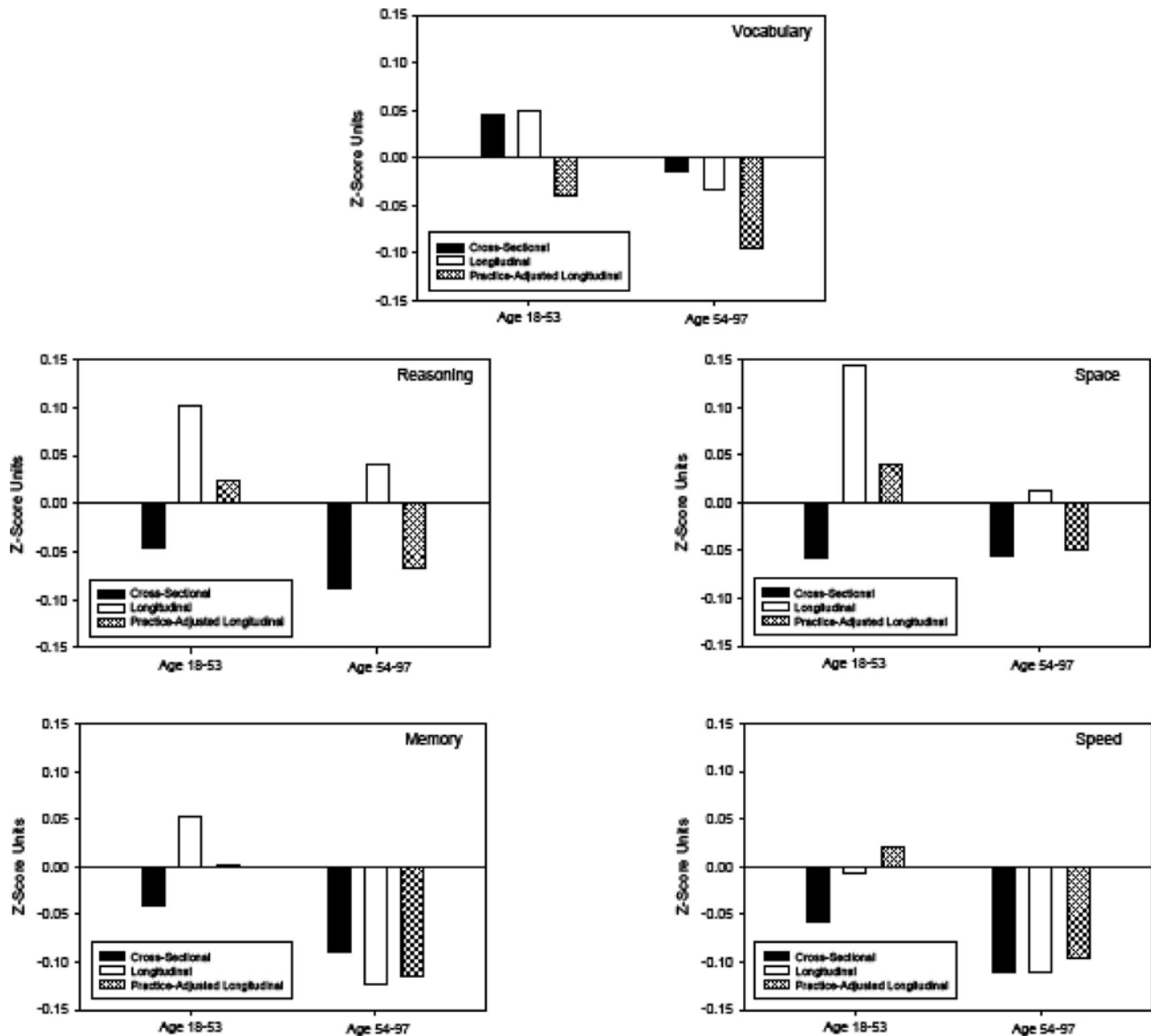


Figure 5. Estimates of cross-sectional differences (over 2.5 years), observed longitudinal change and practice-adjusted longitudinal change for five abilities in adults under and over age 55.

(black bar) and the longitudinal age change (white bar) for all cognitive abilities except vocabulary. The cross-sectional trends were negative and the longitudinal trends were positive for reasoning, space, and memory abilities, and the cross-sectional trends were large and negative whereas the longitudinal trends small and negative for the composite speed variable. After adjusting for practice, most of the longitudinal gains were reduced, and the adjusted longitudinal changes were negative for the composite vocabulary measure. These patterns are similar to those reported by Salthouse (2009a) based on data from a subset of the young adults from the current sample, and slightly different analytical procedures.

The pattern with older adults was somewhat different than that with young adults as there was little discrepancy between the cross-sectional and longitudinal age trends for the composite vo-

cabulary, memory, and speed variables. With these measures the age trends above 55 years of age were similar in cross-sectional and longitudinal comparisons. Furthermore, there was little shift in the age relation in these measures after controlling for practice, with the exception of vocabulary in which the practice-adjusted change was negative. A discrepancy was apparent in the reasoning and space composites, with negative cross-sectional differences and positive longitudinal changes, but after adjusting for practice the longitudinal changes were very similar in direction and magnitude to the cross-sectional differences.

Discussion

The results of these analyses indicate that the interpretation of different age trends in cross-sectional and longitudinal compari-

sons is complicated, with somewhat different patterns across cognitive abilities and at different ages. With some combinations of age and ability the discrepancy appears to be completely attributable to practice effects distorting longitudinal comparisons, as in the reasoning and space abilities in the older adults. In other combinations the discrepancy appears to be partially attributable to practice effects operating in longitudinal comparisons, as in the reasoning, space, and memory abilities in the young adults. However, it is important to note that with certain combinations there is no discrepancy between the cross-sectional and longitudinal age trends. Moreover, in some of those cases there was no effect of practice on the longitudinal age trends, as in the memory and speed abilities in the older adults, and with some combinations adjusting for practice effects resulted in more negative age relations, as in the vocabulary composite in the young adults. Because the effects of practice appear to differ as a function of age, these results imply that not only might different neurobiological substrates be expected for different types of cognitive variables, but also for the same variables at different ages.

Estimates of the practice effects in these analyses are only available at the group level, but with only one exception there were strong negative relations with age. That is, correlations between the mean age and the estimated practice effects in the seven age decades were $-.66$ for vocabulary, $-.09$ for reasoning, $-.89$ for space, $-.86$ for memory, and $-.80$ for speed. These age differences in the benefits of prior experience may be another manifestation of cognitive declines, in that the benefits were smallest at ages where the average level of cognitive performance was lowest. Consistent with this interpretation, Salthouse and Tucker-Drob (2008) found that the short-term practice gains in every cognitive test were positively correlated with an estimate of general cognitive ability.

Some of the practice effects may be attributable to memory of specific items because the same test versions were administered on both occasions. This may be a major factor contributing to the practice effects in the memory tasks. However, Salthouse and Tucker-Drob (2008) found that short-term practice effects, and age differences in the magnitude of practice effects, are apparent even when different test versions are administered on successive sessions. Some of these practice effects, such as in the space tests, may be associated with the discovery of an appropriate strategy for solving these novel types of problems. Although relatively little is currently known about the factors responsible for practice effects in longitudinal research, the results of the current study suggest that these effects are often larger at younger ages, and thus are likely contributing to the differential age trends found in cross-sectional and longitudinal comparisons of neurocognitive aging.

A limitation of existing research investigating practice effects is that the analyses are conducted at the group level. An important goal for future research is to obtain sensitive measures of practice effects, as well as cohort or generational effects, at the level of the individual (cf. Salthouse, 2009a). Exploratory analyses with the current data were conducted to attempt to derive practice estimates for each individual by comparing T1 and T2 scores with age-specific values from the total sample and the once-tested sample, but the resulting values were very noisy. Nevertheless, efforts of this type should continue to allow more direct investigation of the relative contributions of maturational and nonmaturational influences on cross-sectional and longitudinal age trends.

To conclude, although the available estimates of practice effects are still quite crude, it appears that influences associated with prior test performance in longitudinal comparisons are likely responsible for some, but not all, of the discrepancy between cross-sectional and longitudinal age trends in measures of neurocognitive functioning. However, much remains to be learned about this phenomenon because the practice effects vary according to age and according to the particular cognitive ability examined, and the reasons for these variations are not yet understood.

References

- Algeri, S., Biagini, L., Manfredi, A., & Pitsikas, N. (1991). Age-related ability of rats kept on a life-long hypocaloric diet in a spatial memory test: Longitudinal observations. *Neurobiology of Aging, 12*, 277–282.
- Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). *Life-span developmental psychology: Introduction to research methods*. Hillsdale, NJ: Erlbaum.
- Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1997). *Differential Aptitude Test*. San Antonio, TX: The Psychological Corporation.
- Brayne, C., Spiegelhalter, D. J., Dufouil, C., Chi, L. Y., Dening, T. R., Paykel, E. S., et al. (1999). Estimating the true extent of cognitive decline in the old. *Journal of the American Geriatrics Society, 47*, 1283–1289.
- Caprioli, A., Ghirardi, O., Giuliani, A., Ramacci, M. T., & Angelucci, L. (1991). Spatial learning and memory in the radial maze: A longitudinal study in rats from 4 to 25 months of age. *Neurobiology of Aging, 12*, 650–607.
- Christensen, H., Mackinnon, A., Jorm, A. F., Korten, A., Jacomb, P., Hofer, S. M., et al. (2004). The Canberra Longitudinal Study: Design, aims, methodology, outcomes and recent empirical investigations. *Aging, Neuropsychology and Cognition, 11*, 169–195.
- Deary, I. J., & Der, G. (2005). Reaction time, age, and cognitive ability: Longitudinal findings from age 16 to 63 years in representative population samples. *Aging, Neuropsychology and Cognition, 12*, 187–215.
- Dellu, F., Mayo, W., Vallee, M., Le Moal, M., & Simon, H. (1997). Facilitation of cognitive performance in aged rats by past experience depends on the type of information processing involved: A combined cross-sectional and longitudinal study. *Neurobiology of Learning and Memory, 67*, 121–128.
- Du, A.-T., Schuff, N., Chao, L. L., Kornak, J., Jagust, W. J., Kramer, J. H., et al. (2006). Age effects on atrophy rates of entorhinal cortex and hippocampus. *Neurobiology of Aging, 27*, 733–740.
- Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- Ferrer, E., Salthouse, T. A., McArdle, J. J., Stewart, W. F., & Schwartz, B. S. (2005). Multivariate modeling of age and retest in longitudinal studies of cognitive abilities. *Psychology and Aging, 20*, 412–422.
- Ferrer, E., Salthouse, T. A., Stewart, W. F., & Schwartz, B. S. (2004). Modeling age and retest processes in longitudinal studies of cognitive abilities. *Psychology and Aging, 19*, 243–259.
- Finkel, D., Reynolds, C. A., McArdle, J. J., & Pedersen, N. L. (2005). The longitudinal relationship between processing speed and cognitive ability: Genetic and environmental influences. *Behavior Genetics, 35*, 535–549.
- Fotinos, A. F., Snyder, A. Z., Gitton, L. E., Morris, J. C., & Buckner, R. L. (2005). Normative estimates of cross-sectional and longitudinal brain volume decline in aging and AD. *Neurology, 64*, 1032–1039.
- Ghisletta, P., & de Ribaupierre, A. (2005). A dynamic investigation of cognitive dedifferentiation with control for retest: Evidence from the Swiss interdisciplinary longitudinal study on the oldest old. *Psychology and Aging, 20*, 671–682.
- Hultsch, D. F., Hertzog, C., Dixon, R. A., & Small, B. J. (1998). *Memory change in the aged*. New York: Cambridge University Press.

- Kennison, R. F., & Zelinski, E. M. (2005). Estimating age change in list recall in Asset and Health Dynamics of the Oldest-Old: The effects of attrition bias and missing data treatment. *Psychology and Aging, 20*, 460–475.
- Liu, R. S. N., Lemieux, L., Bell, G. S., Sisodiya, S. M., Shorvon, S. D., Sander, J. W. A. S., et al. (2003). A longitudinal study of brain morphometrics using quantitative magnetic resonance imaging and difference image analysis. *NeuroImage, 20*, 22–33.
- Lovden, M., Ghisletta, P., & Lindenberger, U. (2004). Cognition in the Berlin Aging Study (BASE): The first 10 years. *Aging, Neuropsychology and Cognition, 11*, 104–133.
- Markowska, A. L., & Savonenko, A. V. (2002). Protective effect of practice on cognition during aging: Implications for predictive characteristics of performance and efficacy of practice. *Neurobiology of Learning and Memory, 78*, 294–320.
- McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology, 38*, 115–142.
- Owens, W. A. (1953). Age and mental abilities: A longitudinal study. *Genetic Psychology Monographs, 48*, 3–54.
- Prins, N. D., van Dijk, E. J., den Heijer, T., Vermeer, S. E., Jolles, J., Koudstaal, P. J., et al. (2005). Cerebral small-vessel disease and decline in information processing speed, executive function and memory. *Brain, 128*, 2034–2041.
- Rabbitt, P., Diggle, P., Holland, F., & McInnes, L. (2004). Practice and drop-out effects during a 17-year longitudinal study of cognitive aging. *Journal of Gerontology: Psychological Science, 59B*, P84–P97.
- Rabbitt, P., Diggle, P., Smith, D., Holland, F., & McInnes, L. (2001). Identifying and separating the effects of practice and of cognitive ageing during a large longitudinal study of elderly community residents. *Neuropsychologia, 39*, 532–543.
- Raven, J. (1962). *Advanced Progressive Matrices, Set II*. London: H. K. Lewis.
- Raz, N., Lindenberger, U., Rodrigue, 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.
- Rodgers, W. L., Ofstedal, M. B., & Herzog, A. R. (2003). Trends in scores on tests of cognitive ability in the elderly U.S. population, 1993–2000. *Journal of Gerontology: Social Science, 58B*, S338–S346.
- Ronnlund, M., & Nilsson, L.-G. (2006). Adult life-span patterns in WAIS-R Block Design performance: Cross-sectional versus longitudinal age gradients and relations to demographic factors. *Intelligence, 34*, 63–78.
- Ronnlund, M., Nyberg, L., Backman, L., & Nilsson, L.-G. (2005). Stability, growth, and decline in adult life span development of declarative memory: Cross-sectional and longitudinal data from a population-based study. *Psychology and Aging, 20*, 3–18.
- Salthouse, T. A. (2004). Localizing age-related individual differences in a hierarchical structure. *Intelligence, 32*, 541–561.
- Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology, 19*, 532–545.
- Salthouse, T. A. (2007). Implications of within-person variability in cognitive and neuropsychological functioning on the interpretation of change. *Neuropsychology, 21*, 401–411.
- Salthouse, T. A. (2009a). When does age-related cognitive decline begin? *Neurobiology of Aging, 30*, 507–514.
- Salthouse, T. A. (2009b). Decomposing age correlations on neuropsychological and cognitive variables. *Journal of the International Neuropsychological Society, 15*, 650–661.
- Salthouse, T. A. (2010). *Major issues in cognitive aging*. New York: Oxford University Press.
- Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology, 27*, 763–776.
- Salthouse, T. A., Fristoe, N., & Rhee, S. H. (1996). How localized are age-related effects on neuropsychological measures? *Neuropsychology, 10*, 272–285.
- Salthouse, T. A., Pink, J. E., & Tucker-Drob, E. M. (2008). Contextual analysis of fluid intelligence. *Intelligence, 36*, 464–486.
- Salthouse, T. A., Schroeder, D. H., & Ferrer, E. (2004). Estimating retest effects in longitudinal assessments of cognitive functioning in adults between 18 and 60 years of age. *Developmental Psychology, 40*, 813–822.
- Salthouse, T. A., & Tucker-Drob, E. M. (2008). Implications of short-term retest effects for the interpretation of longitudinal change. *Neuropsychology, 22*, 800–811.
- Scahill, R. I., Frost, C., Jenkins, R., Whitwell, J. L., Rossor, M. N., & Fox, N. C. (2003). A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. *Archives of Neurology, 60*, 989–994.
- Schaie, K. W. (1989). Perceptual speed in adulthood: Cross-sectional and longitudinal studies. *Psychology and Aging, 4*, 443–453.
- Schaie, K. W. (2005). *Developmental influences on adult intelligence: The Seattle Longitudinal Study*. New York: Oxford University Press.
- Schaie, K. W., Labouvie, G. V., & Barrett, T. J. (1973). Selective attrition effects in a fourteen-year study of adult intelligence. *Journal of Gerontology, 28*, 328–334.
- Singer, T., Verhaegen, 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.
- Taki, Y., Kinomura, S., Sato, K., Goto, R., Kawashima, R., & Fukuda, H. (2009). A longitudinal study of gray matter volume decline with age and modifying factors. *Neurobiology of Aging*, doi:10.1016/j.neurobiolaging.2009.05.003
- Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale: Third edition*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1997b). *Wechsler Memory Scale: Third edition*. San Antonio, TX: The Psychological Corporation.
- Wilson, R. S., Beckett, L. A., Barnes, L. L., Schneider, J. A., Bach, J., Evans, D. A., et al. (2002). Individual differences in rates of change in cognitive abilities of older persons. *Psychology and Aging, 17*, 179–193.
- Wilson, R. S., Li, Y., Bienias, J. L., & Bennett, D. A. (2006). Cognitive decline in old age: Separating retest effects from the effects of growing older. *Psychology and Aging, 21*, 774–789.
- Woodcock, R. W., & Johnson, M. B. (1990). *Woodcock-Johnson Psycho-Educational Battery: Revised*. Allen, TX: DLM.
- Zachary, R. A. (1986). *Shipley Institute of Living Scale: Revised*. Los Angeles: Western Psychological Services.

Appendix

Description of Reference Variables and Sources of Tasks

Variable	Description	Reliability			Source
		1	2	3	
Reasoning					
Matrix reasoning	Determine which pattern best completes the missing cell in a matrix	.81	.76	.77	Raven (1962)
Shipley abstraction	Determine the words or numbers that are the best continuation of a sequence	.86	.75	.83	Zachary (1986)
Letter sets	Identify which of five groups of letters is different from the others	.79	.65	.66	Ekstrom et al. (1976)
Space					
Spatial relations	Determine the correspondence between a 3-D figure and alternative 2-D figures	.89	.65	.84	Bennett et al. (1997)
Paper folding	Determine the pattern of holes that would result from a sequence of folds and a punch through folded paper	.75	.68	.71	Ekstrom et al. (1976)
Form boards	Determine which combinations of shapes are needed to fill a large shape	.88	.64	.74	Ekstrom et al. (1976)
Memory					
Logical memory	Number of idea units recalled across three stories	.72	.63	.65	Wechsler (1997b)
Word recall	Number of words recalled across trials 1 to 4 of a word list	.90	.71	.69	Wechsler (1997b)
Paired associates	Number of response terms recalled when presented with a stimulus item	.80	.68	.67	Salthouse et al. (1996)
Speed					
Digit symbol	Use a code table to write the correct symbol below each digit	NA	.86	.83	Wechsler (1997a)
Letter comparison	Same/different comparison of pairs of letter strings	.84	.65	.75	Salthouse & Babcock (1991)
Pattern comparison	Same/different comparison of pairs of line patterns	.89	.71	.70	Salthouse & Babcock (1991)
Vocabulary					
Vocabulary	Provide the definitions of words	.89	.77	.80	Wechsler (1997a)
Picture vocabulary	Name designated objects	.85	.76	.82	Woodcock & Johnson (1990)
Synonym vocabulary	Select the best synonym of the target word	.84	.68	.86	Salthouse (1993)
Antonym vocabulary	Select the best antonym of the target word	.83	.61	.74	Salthouse (1993)

Note. Reliability 1 is coefficient alpha (internal consistency reliability), Reliability 2 is the average correlation with scores on parallel tests at the same occasion (alternate forms reliability), and Reliability 3 is the stability correlation across the two occasions (test-retest reliability). NA means that the estimate was not available because there was only one score in the test.

Received November 28, 2009
 Revision received January 13, 2010
 Accepted January 15, 2010 ■