When does age-related cognitive decline begin?

Timothy A. Salthouse *

Department of Psychology, University of Virginia, Charlottesville, VA 22904-4400, United States

Received 17 April 2008; received in revised form 20 August 2008; accepted 12 September 2008

Available online 20 February 2009

Abstract

Cross-sectional comparisons have consistently revealed that increased age is associated with lower levels of cognitive performance, even in the range from 18 to 60 years of age. However, the validity of cross-sectional comparisons of cognitive functioning in young and middle-aged adults has been questioned because of the discrepant age trends found in longitudinal and cross-sectional analyses. The results of the current project suggest that a major factor contributing to the discrepancy is the masking of age-related declines in longitudinal comparisons by large positive effects associated with prior test experience. Results from three methods of estimating retest effects in this project, together with results from studies comparing non-human animals raised in constant environments and from studies examining neurobiological variables not susceptible to retest effects, converge on a conclusion that some aspects of age-related cognitive decline begin in healthy educated adults when they are in their 20s and 30s.

© 2009 Elsevier Inc. All rights reserved.

Keywords: Cognitive aging; Early adulthood; Normal aging

Although there have been many reports over the last 100 years of age-related differences in cognitive functioning, there is still considerable controversy about the age at which cognitive decline begins. This lack of consensus is unfortunate because the question is important for both practical and theoretical reasons. For example, the age at which cognitive decline begins is relevant to the optimum time to implement interventions designed to prevent or reverse age-related declines. Many interventions currently target adults 60 years of age and older. However, if people start to decline when they are in their 20s and 30s, a large amount of change will likely have already occurred by the time they are in their 60s and 70s. This may affect the likelihood that interventions at that age will be successful because the changes might have accumulated to such an extent that they may be difficult to overcome.

The question of when decline begins is also relevant to the theoretical investigation of potential causes of declines in cognitive functioning because declines that begin early are unlikely to be attributable to conditions specific to later life, such as menopause, retirement from paid employment, or certain age-related diseases. The answer to the question of when decline begins may also indicate which period in adulthood is likely to be most informative for learning about causes of age-related cognitive decline because, for example, studies restricted to samples of older adults might have limited value for discovering the causes of a phenomenon that originated decades earlier.

One type of evidence suggesting that age-related cognitive declines begin relatively early in adulthood are the age trends in a variety of neurobiological variables that can be assumed to be related to cognitive functioning. Among the variables that have been found to exhibit nearly continuous age-related declines in cross-sectional comparisons beginning when adults are in their 20s are measures of regional brain volume (Allen et al., 2005; Fotenos et al., 2005; Kruggel, 2006; Pieperhoff et al., 2008; Sowell et al., 2003), myelin integrity (Hsu et al., 2008; Sullivan and Pfefferbaum, 2006), cortical thickness (Magnotta et al., 1999; Salat et al., 2004), serotonin receptor binding (Sheline et al., 2002), striatal dopamine binding (Erixon-Lindroth et al., 2005; Volkow et al., 2000), accumulation of neurofibrillary tangles (Del Tredici and Braak, 2008), and concentrations of various brain metabolites (Kadota et al., 2001).
Furthermore, cross-sectional declines in comparisons of cognitive functioning based on samples of 250 or more adults across a wide age range have been reported since the 1930s (Jones and Conrad, 1933), and have been described in numerous recent publications (Salthouse, 1998, 2005; Salthouse et al., 2003; Schroeder and Salthouse, 2004; Schaie, 2005). In virtually every case, the age trends in these studies have revealed nearly monotonic declines in average level of cognitive performance starting in early adulthood.

It might appear on basis of these well-replicated results with neurobiological and cognitive variables that there is a simple answer to the question of when cognitive decline begins. That is, because cross-sectional age comparisons have consistently revealed nearly continuous age-related decreases in presumably relevant neurobiological variables and in various measures of cognitive performance that appear to begin when adults are in their 20s or early 30s, one might conclude that cognitive decline begins shortly after individuals reach maturity. However, in striking contrast to these empirical results are numerous assertions that cognitive decline begins late in life:

“Cognitive decline may begin after midlife, but most often occurs at higher ages (70 or higher).” (Aartsen et al., 2002).

“...relatively little decline in performance occurs until people are about 50 years old.” (Albert & Heaton, 1988).

“...cognitive abilities generally remain stable throughout adult life until around age sixty.” (Plassman et al., 1995).

“...no or little drop in performance before age 55...” (Ronnlund et al., 2005).

“...most abilities tend to peak in early adulthood, plateau until the late fifties or sixties, and then show decline, initially at a slow pace, but accelerating as the late seventies are reached.” (Schaie, 1989).

A dramatic discrepancy therefore exists between a substantial body of empirical results on one hand, and frequent claims about the time course of cognitive aging on the other hand. Because one cannot hope to explain a phenomenon until its nature, including its trajectory, is accurately described, it is essential to understand the reasons for this discrepancy.

Some of the differences between the evidence just mentioned and the cited assertions may be attributable to variations in how the same findings are interpreted, or to emphases on different types of cognitive variables. However, it is likely that a major reason for the discrepancy is that different patterns of age-cognition relations have been found with longitudinal, or within-person, comparisons, and with cross-sectional, or between-person, comparisons. One of the first reports of a longitudinal comparison with cognitive variables was described in a 1928 book (Thordike et al., 1928). Although other researchers at about the same time reported cross-sectional declines between 18 and 50 years of age on the Army Alpha test, these authors described a study in which the scores for people between 16 and 45 years of age increased over a 5–9-year interval. Rather than revealing decline, therefore, these results suggested that there were improvements in cognitive functioning with increased age when the comparisons were based on observations of the same people at different ages. Subsequent longitudinal studies have replicated the finding of relatively preserved, or even enhanced, levels of cognitive functioning with increased age in longitudinal comparisons involving adults up to about 60 years of age.

Fig. 1 illustrates these patterns with cross-sectional and longitudinal age trends on two tests from the Seattle Longitudinal Study (Schaie, 2005). The top two panels illustrate that there are nearly monotonic age-related declines in the cross-sectional comparisons (dotted lines), but that longitudinal comparisons (solid lines) reveal either stable or increasing age trends. The bottom two panels portray the same data in a different format. In these figures the vertical axis corresponds to standard deviation units rather than T-scores, and the bars represent the cross-sectional difference (black bars) or the longitudinal changes (gray bars). In order to maximize comparability with the results of the current project in which the average retest interval was 2.5 years, the 7-year differences and changes in these figures have been converted to a 2.5-year interval by algebraic substitution (i.e., \(X = 2.5 \times (\text{score}/7)\)). Despite the different formats, the upper and lower panels reveal the same pattern of moderately large negative age trends in the cross-sectional comparisons (dotted lines and black bars), but little or no age decline in the longitudinal comparisons (solid lines and gray bars).

It is not surprising that divergent results such as those portrayed in Fig. 1 have led some researchers to the conclusion that little or no cognitive decline occurs before about age 60. However, a critical assumption of this interpretation is that the results of longitudinal comparisons are more accurate or valid than cross-sectional comparisons with respect to “true” age relations, and it is important to consider what might be responsible for the different patterns of results in the two types of comparisons before accepting this assumption. Only after this issue is resolved can a definitive conclusion be reached about when cognitive decline begins because decline may begin late if cross-sectional results are misleading, but decline may begin early if longitudinal results were found to be influenced by a variety of non-maturational determinants in addition to the maturation effects of primary interest.

What is likely the dominant interpretation of the different age trends found in cross-sectional and longitudinal comparisons of cognitive functioning attributes the discrepancy to characteristics other than age confounding cross-sectional comparisons. Kuhlen (1940) may have been the first to describe what are now commonly referred to as cohort effects, which include a variety of influences on cognitive functioning associated with changes in the social and cultural environment, such as quantity and quality of education, nature of health care, etc.

Although the cohort interpretation is widely accepted, it is currently somewhat underspecified. For example, a criti-
cal expectation of the cohort interpretation is that statistical control of cohort-defining variables should eliminate the cross-sectional age trends. Some variables, such as years of education, are easily assessed, but the prediction that cross-sectional age trends would be eliminated after adjusting for these other variables cannot be adequately tested until all of the cohort-relevant variables are identified and measured. Another limitation of the cohort interpretation is that little is currently known about the time course of cohort influences relative to the age range at which cross-sectional age differences are apparent. That is, cohort influences are sometimes referred to as generational effects, as though they occur over intervals of 25 years or more, but they would have to operate over periods as short as 5 or 10 years to account for the cross-sectional age differences found in some cognitive variables.

Another factor that has been mentioned as a possible contributor to the different cross-sectional and longitudinal age trends is retest effects associated with prior testing. Retest effects refer to influences on the difference in performance between the first and a subsequent measurement occasion that are attributable to the previous assessment. That is, the mere fact that an individual has already been evaluated could change his or her performance on a successive measurement occasion, in which case the age trends inferred from longitudinal comparisons may be misleading with respect to “true” age effects. Cross-sectional comparisons do not involve testing the same individuals again, and therefore retest effects could be contributing to the discrepancy between cross-sectional and longitudinal results by distorting the age trends in longitudinal comparisons. Although seldom mentioned in discussions of the discrepancy between cross-sectional and longitudinal age trends, several findings appear more consistent with the retest interpretation than with the cohort interpretation.

First, because non-human laboratory animals are typically raised in nearly constant environments, age comparisons in non-human animals can be assumed to be free of cohort contaminations attributable to changing environments. To the extent that cohort differences distort cross-sectional comparisons, therefore, little or no age differences in cognitive functioning might be expected in comparisons of non-human animals. However, there are numerous reports of cross-sectional age-related declines in measures of memory and cognition in species ranging from non-human primates (Herndon et al., 1997) to fruit flies (Le Bourg, 2004). Second, although relatively few longitudinal studies have been conducted with non-human animals, it is noteworthy that several studies with rats have reported smaller longitudinal age changes than cross-sectional age differences in measures of maze learning (Caprioli et al., 1991; Dellu et al., 1997; Markowska and Savonenko, 2002). Because the different cross-sectional and

Fig. 1. Estimates of cross-sectional differences and longitudinal changes over 7 years in two variables from the Seattle Longitudinal Study. Cross-sectional data from Table 4.2 and longitudinal data from Table 5.1 of Schaie (2005). The figures in the top two panels portray results of cross-sectional (dotted lines) and longitudinal (solid lines) comparisons in T-score units. The figures in the bottom two panels portray the same data as differences or changes over a 2.5-year interval in standard deviation units.
longitudinal age trends cannot be attributed to cohort differences distorting the cross-sectional results in organisms raised in constant environments, the discrepancies in these studies are likely attributable to retest effects distorting the longitudinal comparisons. And third, several studies examining regional brain volume, which is a variable presumably related to cognitive functioning but not susceptible to practice effects, have reported longitudinal age declines that are at least as large as the cross-sectional differences (Fotenos et al., 2005; Raz et al., 2005; Seihill et al., 2003).

Although the results just described are consistent with the interpretation that longitudinal age trends in cognitive functioning are distorted by the presence of retest effects, they are all indirect. The primary prediction from the retest interpretation is that estimates of retest effects should be moderately large and positive, such that they offset any negative effects of aging or maturation. Several methods have been proposed to estimate the magnitude of retest effects, but only a few have been applied to adults under the age of 60, which is the period most relevant to the question of when age-related decline begins. One method of estimating retest effects is based on a comparison of the cognitive performance in samples of people tested once with those tested twice (Ronnlund et al., 2005; Schaie, 2005). This difference, after adjusting for any differences in initial level of performance, has been used as an estimate of the benefit of prior test experience. Most of the estimates derived from this method have been positive, and considerably larger than the annual cross-sectional age differences. A second method of assessing retest effects has relied on variability across research participants in the retest intervals to decompose the observed change into maturational effects and retest effects. Because this latter method requires a special type of longitudinal design in which people vary in the interval between successive assessments, such that there is not a perfect confounding of the increase in age and the increase in test experience, it has only rarely been used. Nevertheless, both McArdle et al. (2002) and Salthouse et al. (2004) found that in adults under the age of 60, the retest estimates derived from this method were positive and moderately large.

Three different methods of estimating retest effects in longitudinal studies were examined in the current project. As just mentioned, two of the methods, comparing performance of people of the same age tested twice with those tested once, and capitalizing on variability in the retest intervals to distinguish maturational and retest components of longitudinal change, have been used in earlier research. A new method relied on a comparison of the magnitude of change in a longitudinal study with the change in a short-term retest study in which the test–retest interval ranged from 1 to 14 days as the primary basis for distinguishing maturation and retest effects. The rationale is that it is very unlikely that maturation influences are operating over such a short interval, and thus the results of short-term retest studies provide a relatively pure estimate of the potential impact of retest influences that can be compared with estimates of longitudinal change. Furthermore, because the retest interval in the longitudinal study varied from 1 to 7 years, the effect of the retest interval on the magnitude of longitudinal change can also be examined to determine the time course of these retest influences.

1. Methods

1.1. Sample

Characteristics of the samples of participants are summarized in Table 1. The research participants were recruited from newspaper advertisements, flyers, and referrals from other participants, and all were tested individually. The cross-sectional sample included the first assessment from the participants in the samples with longitudinal and short-term retest data, plus additional participants from other studies (Salthouse, 2004, 2005). All research participants were between 18 and 60 years of age, and most participants rated their health as “very good” or “excellent”.

Age-adjusted scaled scores on four standardized tests (i.e., WAIS III Vocabulary and Digit Symbol and WMS III Word Recall and Logical Memory) can be used to assess the representativeness of samples. Scaled scores have means of 10 and standard deviations of 3 in the nationally representative normative samples (cf., Wechsler, 1997). The values in Table 1 indicate that the scaled score means in the current samples were about 1/2 to 1 S.D. above the means in the normative sample, but because the standard deviations were close to 3, it can be inferred that the samples had nearly the same magnitude of between-person variability as the nationally representative normative sample.

The retest intervals in the longitudinal study varied across participants, and ranged from 1 to 7 years, with an average of 2.5 years. The number of participants at each retest interval was 122 at 1 year, 295 at 2 years, 186 at 3 years, 88 at 4 years

<table>
<thead>
<tr>
<th>Table 1 Characteristics of samples.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>Age (years)</td>
</tr>
<tr>
<td>% Females</td>
</tr>
<tr>
<td>Education (years)</td>
</tr>
<tr>
<td>Health</td>
</tr>
<tr>
<td>Scaled scores</td>
</tr>
<tr>
<td>Vocabulary</td>
</tr>
<tr>
<td>Digit Symbol</td>
</tr>
<tr>
<td>Logical Memory</td>
</tr>
<tr>
<td>Word Recall</td>
</tr>
<tr>
<td>Retest Interval</td>
</tr>
</tbody>
</table>

Note: Education is in years of formal education completed, and health is a self-rating of health on a scale ranging from 1 for excellent to 5 for poor. Vocabulary is the raw score on the Wechsler Adult Intelligence Scale III (Wechsler, 1997) Vocabulary test. Retest interval is in years for the longitudinal sample, and in days for the short-term retest sample. N.A. means that the variable is not available in that sample.
and 36 at 5 or more years. There was no correlation between age and retest interval (i.e., \( r = .02 \)). The retest intervals in the short-term retest study ranged from about 1 to 14 days.

1.2. Variables

Nearly all of the participants performed the complete battery of 12 tests, with identical versions of the tests administered on the second test session in the longitudinal and short-term retest studies. Three different types of tests were used to assess the cognitive abilities of inductive reasoning, spatial visualization, episodic memory, and perceptual speed. The cognitive tests, and their reliabilities, are listed in the Supplementary materials and have been described in previous reports (Salthouse, 2004; Salthouse et al., 2003), as have results of confirmatory factor analyses establishing that the variables can be assumed to represent four distinct cognitive abilities.

2. Results

All of the cognitive variables were converted to z-scores by subtracting the score from the total sample mean at the first assessment, and dividing the difference by the standard deviation (S.D.). Fig. 2 portrays the cross-sectional age trends for the 12 variables, where it can be seen that every variable had a linear relation with age. Three variables, matrix reasoning, form boards, and pattern comparison, also had significant quadratic trends, but in each case the quadratic trend was associated with less than 0.2% of the variance. The performance difference from age 18 to age 60 was about 1 S.D. for the speed and spatial visualization variables, and was between .6 and .7 S.D. for the reasoning and memory variables.

The simple correlations of the variables with age ranged from −.07 to −.41, with a median of −.26. The correlations were slightly more negative (i.e., median of −.36) after control of years of education, self-rated health, number of medications taken per week, report of current or past neurological treatment, reported loss of consciousness for more than 5 min, a measure of near-vision visual acuity, and self-reported measures of depression and anxiety (the actual correlations are provided in Supplementary materials). These results therefore imply that the cross-sectional age-cognition relations in this project are not induced by age-related variations in these particular characteristics. Analyses were also conducted to determine the age group with the highest mean score, and the next older age group in which the mean was significantly (\( p < .01 \)) lower. The peak age across the 12 variables ranged from 22 to 27, and the next older age at which the mean was significantly different from the peak age ranged from 27 to 42 (the details for each variable are in the Supplementary materials).

Cross-sectional and longitudinal results for each cognitive variable are displayed in Fig. 3. The longitudinal results represent the change over the average interval of 2.5 years. The cross-sectional differences correspond to the slope of the regression equation relating the cognitive variable to
Fig. 3. Estimates of cross-sectional differences and longitudinal changes over 2.5 years, of the short-term retest effect, and of retest effects derived from two analytical methods for 12 cognitive variables. The vertical axis is in standard deviation units.

age, multiplied by 2.5 to have a comparable interval for the cross-sectional differences and the longitudinal changes. Inspection of the figure reveals that most of the variables have a pattern similar to that in earlier studies with negative cross-sectional differences, and either stable or positive longitudinal changes. If anything, the discrepancy between the cross-sectional and longitudinal age trends in this study was larger than that in the Seattle Longitudinal Study (Schaie, 2005). That is, the average cross-sectional and longitudinal differences in the bottom panels of Fig. 1 were −.06 and .01 S.D. units, respectively, and the medians for the reasoning and spatial visualization variables in Fig. 3 were −.05 for the cross-sectional difference and .08 for the longitudinal change.

Fig. 3 also portrays the changes from the short-term retest study in which the average retest interval was about 1 week. It can be seen that all of the short-term retest changes were positive, and considerably larger than the negative cross-sectional age differences. Finally, Fig. 3 portrays estimates of the retest effects from the difference between scores of participants taking the test for the second time compared to participants taking it for the first time, and from the mixed effects regression analyses based on the variable retest intervals (the estimates were derived from procedures very similar to those used in earlier studies, and details are provided in the Supplementary materials). Although there is clearly some variation in the absolute magnitude of the estimates, it is important to note that nearly all of the estimated retest effects were positive, and substantially larger than the cross-sectional differences.

A series of independent groups t-tests was conducted comparing the short-term retest changes with the longitudinal changes for the 122 participants who had a 1-year interval between assessments. Most of the variables had significantly more positive changes in the short-term retest group than in the 1-year longitudinal group. The only exceptions were one speed test (Letter Comparison), one reasoning test (Letter Sets), and the three spatial tests (Spatial Relations, Paper Folding, and Form Boards). These results indicate that for most of the variables the benefits of prior test experience diminished over the 1 week to 1-year interval after the initial assessment.

A final set of analyses examined the relation between the length of the interval between the two assessments and the magnitude of longitudinal change in the longitudinal sample. The only variables with significant interval effects were the Digit Symbol, Pattern Comparison, Word Recall, and Paper Folding variables. In each case the change values were smaller as the interval increased, with slopes ranging from −.12 to −.16 z-score units per year. For these specific variables, therefore, the change can be predicted to reach zero at retest intervals of 3.6 years (Digit Symbol), 1.6 years (Pattern Comparison), 2.5 years (Word Recall), and 2.4 years (Paper Folding). The lack of significant interval effects for the other variables precludes meaningful estimates of the decay of the retest effects for those variables.
3. Discussion

Fig. 2 reveals that there were significant negative relations between age and several different types of cognitive measures for healthy educated adults ranging from 18 to 60 years of age. Furthermore, additional analyses revealed that the age relations were not attributable to a variety of plausible confounding variables. These results, together with similar findings in many other studies, clearly establish the existence of cross-sectional age-related declines for many cognitive variables prior to age 60. Characteristics related to cohort influences may still be contributing to some of the cross-sectional differences, but these characteristics need to be identified and measured in order for this interpretation to be directly investigated.

It is apparent in Fig. 3 that many cognitive variables exhibit the typical pattern of negative cross-sectional age trends and either stable or positive longitudinal age trends. The unique feature of the current project is that three estimates of retest effects were also available for every variable. Almost all of these estimates were positive, and generally much larger in magnitude than the cross-sectional age differences, which suggest that longitudinal comparisons probably underestimate age-related change for many cognitive variables.

If one assumes that longitudinal changes reflect a mixture of maturation effects and retest effects, a crude method of adjusting the longitudinal changes for retest effects is to subtract the retest estimates from the observed longitudinal changes. Inspection of Fig. 3 reveals that this type of adjustment would dramatically alter the magnitude, and direction, of the longitudinal age trends, and make them much more similar to the cross-sectional patterns. This subtraction method is undoubtedly too simplistic, but only if the retest effects were very small would they not be expected to influence the second assessment in a longitudinal study, and this was not the case for most of the variables in this project.

Inspection of Fig. 3 reveals that there was variability in the magnitude of the retest estimates across variables, and across analytical methods for the same variables. The reasons for this variation are not yet clear, but the results signify the importance of being cautious in basing conclusions on a single variable, or a single analytical method. Despite the differences in the magnitudes of the retest effects, it is important to emphasize that most of the estimates were positive, indicating that the retest phenomenon is robust across different analytical methods, and sets of assumptions.

To summarize, the results of the current project, together with results from research on non-human animals and on neurobiological variables, suggest that age-related cognitive decline begins relatively early in adulthood, but that it may not be detected in longitudinal comparisons until effects of prior test experience are taken into consideration. Not all aspects of cognitive functioning exhibit early age-related declines because measures based on accumulated knowledge, such as performance on tests of vocabulary or general information, are consistently found to increase until at least age 60. However, only those variables exhibiting negative age-related differences in cross-sectional comparisons are directly relevant to the question of when age-related cognitive decline begins.

There is some evidence that the magnitude of age-related decline accelerates at older ages. To illustrate, a sample of about 800 adults between 61 and 96 years of age in my laboratory had cross-sectional slopes of about −.04 to −.05 S.D. units per year compared to the slopes of −.02 to −.03 S.D. units per year for adults under age 60. In absolute units, the decline in speed variables was about twice as great in this age range compared to adults under age 60, and the decline in the memory variables was nearly four times greater. What is not yet known is whether these quantitatively different age trends reflect changes in the same set of influences, or the operation of qualitatively different types of influences. However, what does appear clear is that several different types of results converge on the conclusion that age-related cognitive decline begins relatively early in adulthood, and certainly before age 60 in healthy educated adults.

Finally, although the results of this project suggest that at least some of the differences in the age trends in cross-sectional and longitudinal comparisons are attributable to retest effects distorting longitudinal comparisons, the results should not be interpreted as implying that longitudinal comparisons are not meaningful or valuable. In fact, quite the opposite is true because only with longitudinal data can one examine within-individual changes distinct from between-person differences. Instead, the major point is that merely because the changes are observed within the same individual does not mean that they only reflect aspects of maturation. Strengths and weaknesses of both cross-sectional and longitudinal data therefore need to be considered when reaching conclusions about age trends in cognitive functioning.

Conflict of interest statement

The author has no financial or other conflicts related to this research.

Acknowledgements

Supported by Grant R37AG024270 from the National Institute on Aging.

Institutional Review Board approval: The research described in this report was conducted with approval of the Institutional Review Board at The University of Virginia.

Appendix A. Supplementary data

References


