Within-cohort age differences in cognitive functioning

Timothy A. Salthouse
Department of Psychology
University of Virginia
Charlottesville, VA 22904
Email: Salthouse@virginia.edu

Abstract

It is widely accepted that the level of cognitive functioning can be influenced by characteristics of the environment that change over time. Many developmental researchers have referred to these influences as cohort effects, and have used year of birth as the basis for determining cohort membership. Furthermore, age-related differences in cognitive functioning are sometimes assumed to be primarily attributable to cohort differences, which implies that differences across different birth cohorts should be much larger than differences within the same birth cohort. Comparisons of the level of performance in five cognitive abilities in different people tested at different ages in different years revealed that within-cohort age differences were often as large as between-cohort age differences. These results lead to questions about the practice of relying on birth cohort to represent influences on cognitive functioning associated with temporal shifts in characteristics of the environment.
There is convincing evidence that historical changes in the physical and social environment can affect the level of different types of cognitive functioning. The phenomenon that cognitive test scores increase across generations in comparisons of people of the same age who are tested at different times is now so well established that it has been termed the Flynn Effect, after one of the first researchers to document its existence (Flynn, 1984, 1987).

The factors responsible for historical increases in cognitive performance have not yet been identified, and consequently many developmental researchers have relied on year of birth as a proxy for these time-related influences. The assumption is that birth year can serve as an index of influences operating over historical time in a manner analogous to how chronological age is often used as an index of maturational influences. People born within the same range of birth years have therefore been categorized as belonging to the same birth cohort, and age-related differences in cognitive functioning are sometimes assumed to be determined as much, or more, by differences associated with cohort membership as by differences associated with chronological age. For example, Schaie, Labouvie and Buech (1973) claimed that “generational differences account for a major share of the variation between different age groups when studied at one point in time (p. 152).” More recently, Schaie (2005) proposed that: “Unless there is independent evidence to suggest that older cohorts performed at the same level as younger cohorts at equivalent ages, it would be most parsimonious to assume, at least in comparisons of adult samples, that cross-sectional age differences represent estimates of cohort differences … (p. 21).”

An implication of the hypothesis that differences in cognitive performance are determined more by when people were born than by their current age is that little or no relation of age to cognitive functioning would be expected in comparisons of people of different ages from the same birth cohort tested at different times. This implication was investigated by Schaie and Strother (1968) and by Schaie et al. (1973) by comparing independent samples of individuals from the same birth years who were tested in different years. Schaie and colleagues suggested that the observed age differences in measures of cognitive functioning in the independent-samples same-cohort comparisons were dramatically different from the patterns with cross-sectional (between-cohort) comparisons, and that they closely resembled the patterns with longitudinal (within-cohort) comparisons. However, discrepancies between the between-cohort and within-cohort age gradients in several of the figures in Schaie et al. (1973) were only evident for adults between about 40 and 60 years of age, and alternative interpretations of the results have been proposed by other researchers. For example, Horn and Donaldson (1976), who referred to the comparison based on independent samples of people of different ages tested at different times as quasi-longitudinal, suggested that the age trends in the within-cohort data were very similar to those in the cross-sectional analyses. Furthermore, secondary analyses of Schaie’s data by Salthouse (1991) revealed that the age relations in independent samples from the same birth cohort were nearly identical to those from conventional between-cohort, cross-sectional comparisons.

More recently, Ronnlund and colleagues (Ronnlund & Nilsson, 2006; Ronnlund et al. 2005) described a project in which the same cognitive tests were administered to independent samples of adults across a wide age range in 1989 and in 1994. Results were reported from a spatial (block design) test, and for factor scores representing episodic memory (based on five memory tests), and semantic memory (based on a vocabulary test, a general knowledge test, and three verbal fluency tests). Comparisons of the tabular data reported in the articles indicate that with the block design variable the within-cohort differences were nearly identical to the between-cohort differences at every age. The patterns with the episodic memory and semantic memory factors were less consistent as the within-cohort contrasts were more positive than the between-cohort contrasts for participants in their 40s and 50s for episodic memory, and for participants in their 40s, 50s and 60s for semantic memory. As in the Schaie data, the patterns of age differences were nearly identical in the within-cohort and between-cohort comparisons for adults older than about 60 years of age.

The preceding review reveals that prior research with independent-sample within-cohort comparisons of age differences in cognitive functioning has been inconsistent, in that the same results have been interpreted differently, and different patterns have been reported within the same project across variables, and at different periods of adulthood for the same variable. Several factors might
account for the inconsistencies in the earlier same-cohort independent samples comparisons. One issue is that samples in different years may not have been comparable, in which case the within-cohort contrasts could have been confounded with sample selectivity. Because even random sampling can result in fluctuations across test years, it is desirable that sample comparability be directly assessed rather than merely assumed.

A second issue regarding prior research is that the boundaries between birth cohorts are arbitrary, and may not correspond to salient physical or social changes. In the studies by Schaie and Ronnlund individuals were grouped into 7-year and 5-year birth cohorts, respectively, presumably because those ranges corresponded to the intervals between test occasions. However, because the relations of birth year to relevant environmental changes are currently unknown, valuable information may be lost by imposing discrete boundaries on a continuous variable.

The current study incorporated these considerations in an examination of between-cohort and within-cohort age trends in measures of cognitive functioning in participants from the Virginia Cognitive Aging Project (VCAP, Salthouse, 2004, 2010a, 2011). The same tests have been administered to new (and returning) samples of adults across a wide age range between 2001 and 2011, and therefore relations of cognitive functioning could be examined as a function of both birth year (within cohort) and test year (between cohort).

The relevant contrasts can be illustrated with the structure in Figure 1, which portrays a subset of the sample in the current study. Notice that people from different birth years (horizontal axis) were tested in different years (vertical axis), with the numbers in the cells representing the age at the time of testing. The highlighted row, corresponding to a traditional cross-sectional comparison, consists of people born in different years and tested in the same year (i.e., 2006). Because the testing extended over multiple years, it was also possible to examine the performance of people born in the same year (i.e., 1946) but tested in different years, as represented by the highlighted column in Figure 1. It is important to note that unlike a longitudinal comparison, which would be represented as a positive diagonal in Figure 1, in these comparisons the people in each cell are different.

Because birth year and test year were both continuous variables, they were used as simultaneous predictors in regression analyses instead of as discrete factors in analyses of variance as in the Ronnlund and Schaie studies. The relevant information from the regression analyses consists of estimates of the relations of birth year (between-cohort) and test year (within-cohort) with the measures of cognitive functioning. Variations in both birth year and test year correspond to differences in age, and therefore the key question is whether there are age-related differences in cognitive functioning in within-cohort comparisons, and if so, what is their magnitude relative to the age-related differences in between-cohort comparisons.

Representativeness of the samples in different years can be evaluated by comparisons of the scores of the sample participants with scores from a nationally representative normative sample. Age-adjusted scaled scores from the Wechsler (WAIS III, Wechsler, 1997a; and WMS III, Wechsler, 1997b) tests were used for this purpose because they are based on a sample selected to match the proportions in the population with respect to major demographic characteristics.

Method

Participants

Characteristics of the participants in each test year are reported in Table 1. Across the entire sample the birth years ranged from 1907 to 1989, chronological age ranged from 18 to 97, and test year ranged from 2001 to 2011. There was a wide range of ages among the participants in each test year, and the correlation between birth year and test year was only .06.

Eighty-one percent of the participants classified themselves as white, 10 percent as black, and the remainder identified themselves as belonging to other ethnicities, or to more than one ethnic group. Sixty-five percent of the participants were females, and the average years of formal education was 15.7.

Between 205 and 913 adults were recruited as new participants each year except 2010. The average scaled scores for tests from the WAIS III battery (Vocabulary and Digit Symbol) and from the WMS III battery (Logical Memory and Word Recall) in each test year are reported in Table 1. The age-adjusted scaled score means for each test in the normative sample are fixed at 10 with a
standard deviation of 3. Most of the averages in Table 1 were between 11 and 13, and therefore the participants in this project can be inferred to be functioning at about the 75th percentile of the nationally representative normative sample. Importantly, the scaled scores varied across test years, as the correlations with test year were -.13 for vocabulary, -.12 for digit symbol, -.09 for logical memory, and -.16 for recall (all p<.01). The samples may therefore have differed in degree of representativeness relative to normative sample, in the direction of somewhat less select samples in more recent years.

**Measures**

Sixteen different tests were used to assess five cognitive abilities. Reasoning was assessed with the Raven’s Progressive Matrices Test (Raven, 1962), the Shipley Abstraction Test (Zachary, 1986), and the Letter Sets Test (Ekstrom et al., 1976). Spatial visualization was assessed with the Spatial Relations Test (Bennett et al., 1997), the Paper Folding Test (Ekstrom et al., 1976), and the Form Boards Test (Ekstrom et al., 1976). Vocabulary was assessed with the Vocabulary test from the WAIS III (Wechsler, 1997a), the Picture Vocabulary test from the Woodcock-Johnson Cognitive Battery (Woodcock & Johnson, 1990), and multiple-choice Synonym and Antonym tests (Salthouse, 1993). Verbal memory was assessed with the immediate Logical Memory and Word Recall Tests from the WMS III (Wechsler, 1997b), and a paired associates test (Salthouse, Fristoe & Rhee, 1996). Perceptual speed was assessed with the Digit Symbol test from the WAIS III (Wechsler, 1997a), and with the Pattern Comparison and Letter Comparison Tests (Salthouse & Babcock, 1991). Previous research has established that the tests were all reliable, and valid in the sense that they had moderate to high loadings on their respective ability factors (e.g., Salthouse, 2004; Salthouse & Ferrer-Caja, 2003; Salthouse et al., 2008).

**Results**

Because composite scores are more reliable and less influenced by test-specific factors than scores from individual tests, all analyses were conducted on composite scores formed by averaging z-scores for the tests representing each ability. Figure 2 portrays the relations averaged across all test years for the composite scores in each ability as a function of birth year (and age). As is typically found, there were negative age relations (positive birth cohort relations) for reasoning, space, memory, and speed abilities, and a positive age relation (negative birth cohort relation) until about age 60 for measures of vocabulary knowledge. Because later birth years (corresponding to younger ages) are associated with higher levels of cognitive performance, the slopes of cognitive performance as a function of birth year were positive in all composite scores except vocabulary, and correspond to differences of about .02 standard deviations per year.

Age differences in between-cohort (birth year) and within-cohort (test year) contrasts were investigated with multiple regression analyses in which birth year and test year were simultaneous predictors of the cognitive ability composite scores. The (unstandardized) regression coefficients for birth year and test year from these analyses are presented in the left side of Table 2. The coefficient for birth year represents the difference per year in the composite score when the variation in test year is controlled at the average test year (i.e., 2005.8).

The coefficient for test year indicates the difference per year in the composite score when the variation in birth year is controlled at the average birth year (i.e., 1954.6). Because increases in test year are associated with increases in age, these coefficients provide an estimate of the age-related differences within the same birth cohort. Note that all of the values were considerably larger than their standard errors, and each was more than twice as large as the estimates of the between-cohort (birth year) age differences.

A second set of analyses included the age-adjusted scaled scores for the vocabulary, digit symbol, logical memory, and word recall tests as covariates to adjust for differences in sample selectivity across test years. The effect of including these variables as covariates is to carry out all comparisons at the average values of the scaled scores (i.e., 12.5 for vocabulary, 11.3 for digit symbol, 11.7 for logical memory, and 12.1 for recall). Results of these analyses are presented in the right panel of Table 2. The estimates of the birth year effects were very similar to the values from the analyses without adjusting for sample selectivity reported in the left panel of Table 2. However, controlling the variation in scaled scores reduced the estimates of the test year effects in all five
cognitive abilities. The estimates were very small for the vocabulary and reasoning composite scores, but the within-cohort estimates of the age differences were approximately the same magnitude as the estimates of the age differences based on different cohorts (birth year) for the space, memory, and speed composite scores.

In order to determine if the birth year and test year relations varied according to birth cohort (and age), the analyses were repeated in separate birth cohort groups corresponding to the birth year boundaries specified in the horizontal axis of Figure 2. Because the standard deviations of the birth year and test year variables were very similar in each group, these comparisons allow the effects of birth year and test year to be evaluated when the range of variation is similar for the two predictor variables. Sample sizes in these groups ranged from 210 to 946, with a median of 597.

The regression coefficients from these analyses, in which the scaled scores were included as additional predictors to control for sample selectivity, are presented in the five panels of Figure 3. The coefficients are estimates of the age differences per year in the between-cohort (birth year) and within-cohort (test year) contrasts within the specified range of birth years. In order to facilitate comparison of the between-cohort and within-cohort effects, the regression coefficients for birth year were multiplied by -1. The effect of this transformation is to express both types of comparisons in the same direction, with negative coefficients corresponding to lower levels of cognitive ability with increased age.

There are two major points to note in Figure 3. The first is that the patterns in different cognitive abilities varied somewhat across birth cohorts (and age groups). The functions were relatively flat for space ability, indicating that the relation between performance and age was nearly constant across successive age/cohort groups. In contrast, the functions were negative for the other abilities, indicating that the age relations were more negative at older ages (and earlier cohorts). The second, and most interesting, point regarding the results in Figure 3 is that each panel had nearly the same magnitude of between-cohort (birth year) and within-cohort (test year) estimates of the age differences. These results imply that age-related differences in cognitive ability at the average test year in between-cohort comparisons were very similar to age-related differences at the average birth year in within-cohort comparisons.

Discussion

The major finding in this report was the discovery of significant within-cohort age differences in five cognitive abilities that were nearly the same magnitude as the age differences in conventional cross-sectional (between-cohort) comparisons. The initial estimates of the within-cohort age differences were considerably more negative than those for the between-cohort age differences. However, because samples in different test years may have differed in selectivity, the estimates after controlling for the scaled scores may provide the best indication of the within-cohort age relations. It is therefore noteworthy that these adjusted within-cohort estimates were nearly the same magnitude as the between-cohort estimates for the space, memory, and speed composite scores.

The variation in scaled scores across test years was assumed to be attributable to differential selectivity of the samples, and thus the scaled scores were used as covariates to equate the samples on these variables. Although it is possible that differences across test years reflect systematic time-lag effects, this seems unlikely because the direction of the effects is opposite of that reported for the Flynn effect. That is, instead of higher cognitive scores in more recent test periods as is typically found, the trends in Table 1 indicate a slight decrease in the average scaled scores from early to later test years. However, the trend is consistent with a broader, and less selective, recruitment of participants in later test years.

The finding that age differences in cognitive functioning are as large in contrasts within the same birth cohort as in contrasts involving people from different birth cohorts implies that birth year is a poor proxy for the changing aspects of the physical and social environment that may influence cognitive performance. Several limitations of operationalizing environmental changes in terms of birth cohort were discussed by Salthouse (1991). Among these were that: “... it is unlikely that individuals will be meaningfully grouped with respect to common experiences when they are classified according to arbitrary temporal boundaries ... not all experiences are presumed to be relevant, and there is no assurance that individuals classified together on the basis of year of birth all share the critical
experiences ... [and] ... because the impact of experiences may not be uniform among individuals born within specified temporal intervals, relying upon birth year to define cohorts also leads to the problem of possible differential effects of critical experiences among individuals treated as equivalent (p. 117).” When these concerns are combined with the current findings that within-cohort age differences are nearly as large as between-cohort age differences, it may no longer be meaningful to refer to age differences in cognitive functioning as cohort differences when cohort is defined exclusively in terms of birth year.

It is important to emphasize that the current results are not inconsistent with the documented time-lag improvements in average scores in cognitive tests (i.e., the Flynn Effect). That is, the Flynn Effect can exist without necessarily being relevant to the interpretation of age differences or age changes in cognitive functioning if these environmental changes affect people of all ages, and not merely people during certain developmental periods. Salthouse (2010b) recently suggested that the historical shifts in relevant aspects of the environment that influence cognitive test scores could be analogous to the effects of inflation on income. That is, when inflation is operating, average income for a given age will be higher in more recent years, but these time-related increases would not necessarily compromise the meaningfulness of cross-sectional age comparisons in any given year. In fact, historical shifts such as the effects of inflation on income would likely result in greater distortions of the maturational trends in longitudinal comparisons than in cross-sectional comparisons. It is not yet clear whether people of all ages benefit equally from time-lag improvements in cognitive functioning, but to the extent that they do, the implications for the interpretation of age relations in cognitive functioning may be similar to those with inflation and income in that the distortion may be more pronounced with longitudinal comparisons rather than cross-sectional comparisons.

Disentangling potential determinants of age-related cognitive change is challenging because there is still limited understanding of what is responsible for the relations between age and cognitive functioning. It is widely recognized that age is not a causal variable, but instead is best conceptualized as a continuum along which causal factors operate. The goal of many developmental researchers has therefore been to identify causal factors that could be used to replace the age variable in explanations. Because the current results suggest that age differences in several cognitive abilities are as large in contrasts within the same birth cohort as in conventional cross-sectional samples involving different birth cohorts, a high priority for future research should be the replacement of birth year as a proxy for time-related environmental influences. Rather than attributing age-related differences to unknown correlates of birth year, a more productive approach would involve specifying and measuring the relevant characteristics, and including them in analyses to evaluate the role of these factors in the age differences and age changes in cognitive functioning. In fact, one could argue that a decompositional strategy such as this is desirable with every presumed cause of human behavior.

The importance of investigating the aspects of the environment responsible for variations in human behavior has also been emphasized by other researchers. For example, Plomin and Daniels (1987) proposed a three-step approach to investigate the nature of environmental influences in behavioral genetics. The three steps were to document differential experiences, document association between differential experiences and differential outcomes, and investigate the extent to which the associations between differential experiences and differential outcomes were causal. If researchers were interested in interventions, an additional step might be to specify and investigate the mechanisms responsible for the environment-behavior associations. Implementing these steps will be difficult, and interpretations can be complicated if different aspects of the environment are relevant for different dimensions of behavior, and at different periods in development. However, environmental influences are too important to be inferred by what is not accounted for by other influences, or with crude proxies that fail to capture the relevant aspects responsible for the behavioral variations of primary interest.
Within-Cohort Age Differences

References


Author Notes

This research was supported by Award Number R37AG024270 from the National Institute on Aging. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institute on Aging or the National Institutes of Health. There are no conflicts of interest.
Table 1
Descriptive characteristics of samples tested for the first time in different years

<table>
<thead>
<tr>
<th>Test Year</th>
<th>N</th>
<th>Age</th>
<th>Birth Year</th>
<th>Females</th>
<th>Health</th>
<th>Education</th>
<th>Vocabulary</th>
<th>Digit Symbol</th>
<th>Logical Memory</th>
<th>Word Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>205</td>
<td>50.7</td>
<td>1950</td>
<td>.65</td>
<td>2.1</td>
<td>16.0</td>
<td>13.3</td>
<td>12.2</td>
<td>12.0</td>
<td>13.0</td>
</tr>
<tr>
<td>2002</td>
<td>269</td>
<td>48.3</td>
<td>1954</td>
<td>.63</td>
<td>2.0</td>
<td>15.9</td>
<td>12.2</td>
<td>11.9</td>
<td>11.5</td>
<td>12.5</td>
</tr>
<tr>
<td>2003</td>
<td>330</td>
<td>47.4</td>
<td>1956</td>
<td>.66</td>
<td>2.0</td>
<td>15.7</td>
<td>12.6</td>
<td>11.9</td>
<td>12.3</td>
<td>13.3</td>
</tr>
<tr>
<td>2004</td>
<td>458</td>
<td>51.3</td>
<td>1953</td>
<td>.63</td>
<td>2.0</td>
<td>15.6</td>
<td>12.9</td>
<td>11.3</td>
<td>12.1</td>
<td>12.4</td>
</tr>
<tr>
<td>2005</td>
<td>670</td>
<td>49.8</td>
<td>1955</td>
<td>.66</td>
<td>2.0</td>
<td>15.4</td>
<td>12.9</td>
<td>11.0</td>
<td>11.7</td>
<td>12.1</td>
</tr>
<tr>
<td>2006</td>
<td>913</td>
<td>51.1</td>
<td>1955</td>
<td>.66</td>
<td>2.5</td>
<td>15.7</td>
<td>12.9</td>
<td>11.4</td>
<td>11.9</td>
<td>12.4</td>
</tr>
<tr>
<td>2007</td>
<td>444</td>
<td>52.3</td>
<td>1955</td>
<td>.62</td>
<td>2.1</td>
<td>15.6</td>
<td>11.9</td>
<td>10.9</td>
<td>11.5</td>
<td>11.5</td>
</tr>
<tr>
<td>2008</td>
<td>255</td>
<td>54.6</td>
<td>1953</td>
<td>.64</td>
<td>2.2</td>
<td>15.5</td>
<td>11.2</td>
<td>10.7</td>
<td>11.1</td>
<td>10.7</td>
</tr>
<tr>
<td>2009</td>
<td>236</td>
<td>57.1</td>
<td>1952</td>
<td>.69</td>
<td>2.1</td>
<td>15.7</td>
<td>11.8</td>
<td>11.1</td>
<td>11.4</td>
<td>11.3</td>
</tr>
<tr>
<td>2011</td>
<td>440</td>
<td>52.5</td>
<td>1958</td>
<td>.68</td>
<td>2.2</td>
<td>15.9</td>
<td>11.7</td>
<td>10.8</td>
<td>11.4</td>
<td>11.3</td>
</tr>
</tbody>
</table>

Note: Health is a self rating on a scale from 1 for "excellent" to 5 for "poor". The vocabulary and digit symbol scaled scores are age-adjusted scores derived from the Wechsler Adult Intelligence Scale III (Wechsler, 1997a), and the Logical Memory and Word Recall scaled scores are age-adjusted scores derived from the Wechsler Memory Scale III (Wechsler, 1997b).
Table 2
Unstandardized regression coefficients (with standard errors) in the total sample for differences per birth year (between cohort) and per test year (within cohort) in composite scores of cognitive functioning

<table>
<thead>
<tr>
<th></th>
<th>Original Analyses</th>
<th></th>
<th>Controlling Scaled Scores</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Birth Year</td>
<td>Test Year</td>
<td>Birth Year</td>
<td>Test Year</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>-.012 (.001)</td>
<td>-.037 (.005)</td>
<td>-.010 (.000)</td>
<td>.006 (.003)</td>
</tr>
<tr>
<td>Reasoning</td>
<td>.020 (.001)</td>
<td>-.051 (.007)</td>
<td>.024 (.001)</td>
<td>-.006 (.005)</td>
</tr>
<tr>
<td>Space</td>
<td>.022 (.001)</td>
<td>-.050 (.005)</td>
<td>.023 (.001)</td>
<td>-.027 (.005)</td>
</tr>
<tr>
<td>Memory</td>
<td>.019 (.001)</td>
<td>-.067 (.005)</td>
<td>.022 (.000)</td>
<td>-.023 (.003)</td>
</tr>
<tr>
<td>Speed</td>
<td>.029 (.001)</td>
<td>-.071 (.004)</td>
<td>.031 (.000)</td>
<td>-.046 (.003)</td>
</tr>
</tbody>
</table>
Figure Captions

Figure 1 – Illustration of the structure of the data in the current project. Although the birth years ranged from 1907 to 1989, only a limited number of birth years are illustrated for the sake of clarity. Each cell in the matrix consists of different people of the specified age, and thus comparisons along a row are between cohort because they involve people of different ages (and thus born in different years) tested in the same year, and comparisons along a column are within cohort because they involve people of different ages (but born in the same year) tested in different years. Numbers in cells correspond to the age of individuals from that birth year in that test year.

Figure 2 – Mean (and standard error) composite scores in five cognitive abilities as a function of birth year (and age).

Figure 3 – Estimates of the differences in composite scores per year in between-cohort (birth year) and within-cohort (test year) comparisons in different birth years (and ages) when scaled scores are used as covariates. Note that the estimates for birth year have been multiplied by -1 to express both sets of coefficients in the same direction (i.e., negative coefficients indicate lower scores with increased age).
Within-Cohort Age Differences