NIH Public Access
Author Manuscript
Dev Psychol. Author manuscript; available in PMC 2014 November 01.

Published in final edited form as:
Dev Psychol. 2013 November ; 49(11): . doi:10.1037/a0032019.

Effects of first occasion test experience on longitudinal cognitive change

Timothy A. Salthouse
Department of Psychology, University of Virginia, Charlottesville, VA 22904-4400

Abstract

Effects of additional test experience on longitudinal change in five cognitive abilities was examined in a sample of healthy adults ranging from 18 to 80 years of age. Participants receiving experience with parallel versions of the cognitive tests on the first occasion had more positive cognitive change an average of 2.5 years later than participants performing only a single version of the tests on the first occasion. Importantly, these test experience effects were similar in adults of different ages, which implies that retest contributions to cognitive change are comparable among healthy adults between 18 and 80 years of age.

It is widely recognized that change in performance on cognitive tests can be influenced by maturational factors and by factors associated with prior testing experience. That is, both time-related developmental processes and reactive effects associated with the initial assessment can contribute to the change in cognitive functioning from a first to a second assessment. Because the interpretations could be quite different if increased age, diseases, or interventions primarily affect the amount of benefit associated with an initial assessment as opposed to time-related change processes, it is important to try to distinguish these influences.

Three major approaches have been used to distinguish contributions of the two components. One approach (e.g., Ferrer et al., 2004; 2005; Rabbitt et al., 2001 Rabbitt et al., 2004) is based on the assumption of different functions for test experience (e.g., step function with coefficients of 0, 1, and 1 representing three successive occasions), and for developmental processes (e.g., linear function with coefficients of 1, 2, and 3 representing three successive occasions). An advantage of this approach is that it can be applied in traditional longitudinal studies with three or more occasions, and it is potentially applicable at the level of individuals and not merely at the level of groups. The primary disadvantage of the dual-function approach is that the assumption of different functions is seldom directly tested, and little is known about possible invariance of the functions across participants, cognitive abilities, and intervals between occasions.

A second approach used to distinguish the two contributions is based on a special type of longitudinal study in which the interval between testing occasions varies across individuals (e.g., McArdle et al., 2002; Salthouse, Schroeder & Ferrer, 2004, Salthouse 2009). In a traditional longitudinal study the interval is approximately the same for all participants, which has the consequence of producing a nearly perfect correlation between the increase in age and the increase in test experience. However, the correlation can be reduced, and the two influences potentially distinguished, if people vary in the intervals between occasions. An advantage of the variable-interval approach is that it can be used with longitudinal data involving only two occasions, but disadvantages are that the variability in the intervals has

There are no conflicts of interest.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

to be incorporated into the research design, and the age-experience correlation will still be high unless the range of longitudinal intervals is quite large.

A third approach used to separate test experience influences from time-related influences is also based on a modification of the typical longitudinal study, in this case by adding a new sample of participants who are tested for the first time when the original longitudinal participants are tested at the second time (e.g., Ronnlund et al. 2005, Ronnlund & Nilsson, 2006; Salthouse, 2009; 2010a; Schaie, 1988). If the groups are truly comparable except for the prior test experience, the difference in their performance provides an estimate of the effects of test experience. The major advantage of this approach is its conceptual simplicity, but disadvantages are that the addition of another sample of participants increases the cost and complexity of the research, and it can be difficult to establish that the groups are equivalent in all respects except for the prior test experience.

Little is known about the relations among the test experience estimates derived from different procedures because apparently only one study has applied more than one approach in the same data, and the estimates based on the different methods were somewhat variable (i.e., Salthouse, 2009). Although analyses with all three types of approaches have generally revealed positive estimates of test experience, the results have been inconsistent with respect to age differences. That is, Salthouse (2010a) found somewhat smaller retest estimates at older ages, little or no age differences in retest effects were reported by Ferrer et al. (2004, 2005), Ronnlund et al. (2005), Ronnlund and Nilsson (2006) and Schaie (1988), and Rabbitt et al. (2004) reported larger retest effects at older ages.

The current project used a novel procedure to investigate the contribution of test experience on cognitive change by capitalizing on the variation in test experience at the first occasion among participants within the Virginia Cognitive Aging Project (VCAP, Salthouse, 2009, 2010a, 2011). All participants described in this report came to the laboratory for three sessions at each of two occasions separated by an average of about 2.5 years. The second occasion was identical for all participants, and involved different versions of the same tests on three successive sessions. However, the participants differed with respect to the nature of their experience on the first occasion. Some of the participants performed three versions of the same tests on successive sessions. The remaining participants performed the reference VCAP tests on the first session in the first occasion, but performed other tests on the second and third sessions in that occasion. These individuals therefore had the same amount of social contact and overall test experience as the individuals with three versions of the same tests, but different tests were performed on the second and third sessions. The other tests were designed to assess fluid intelligence (Salthouse, Pink & Tucker-Drob, 2008), source memory (Siedlecki, Salthouse & Berish, 2005), prospective memory (Salthouse, Berish & Siedlecki, 2004), executive functioning (Salthouse, 2010b; Salthouse, Atkinson & Berish, 2003), cognitive control (Salthouse, Siedlecki & Krueger, 2006), and language production (Rabaglia & Salthouse, 2012). Assignment of participants to one or three versions at the first occasion was determined by the research goals at the time of recruitment, and was not based on level of performance in any of the tests.

The primary question examined in this study was whether additional experience with different versions of the same tests on the first (T1) occasion affected the magnitude of change from the first to the second occasion, and if so, whether these experience effects varied with the age of the research participant.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Method

Sample

The participants in this study were VCAP participants between 18 and 80 years of age who completed at least two occasions, and had Mini Mental State Exam (Folstein et al., 1975) scores above 23 at the second occasion. Participants were recruited by newspaper advertisements, flyers, and referrals from other participants, and they were paid for their participation. The initial (T1) participation occurred between 2001 and 2009, with the second (T2) participation occurring between one and ten years later. The interval between occasions was varied across participants to allow increases in test experience to be distinguished from increases in age (e.g., Salthouse, 2009), and to investigate the relation between longitudinal interval and change at different ages (Salthouse, 2011). The correlation of age with the length of the interval was only −.03 each in the groups with one and three test versions at T1.

Characteristics of the participants with one and three versions of the critical tests at T1 are summarized in Table 1. In most respects the samples with one and three versions at T1 were similar, with the exception of slightly higher (poorer) self ratings of health, lower word recall scaled scores, and a longer average interval between the T1 and T2 occasions in the group with only one T1 version. Analyses of variance on the measures in Table 1 with age decade and number of T1 versions as factors revealed that none of the interactions was significant.

Although VCAP participants can be considered a convenience sample because they were all community volunteers, their representativeness can be evaluated in terms of scores on standardized tests that have been normed in nationally representative samples. Scaled scores from the Vocabulary and Digit Symbol subtests in the Wechsler Adult Intelligence Scale III (Wechsler, 1997a) and from the Logical Memory and Word Recall subtests in the Wechsler Memory Scale III (Wechsler, 1997b) were used for this purpose. The age-adjusted scaled scores have a mean of 10 and a standard deviation of 3 in the normative sample, and thus the values in Table 1 indicate that individuals in the current sample were functioning above average relative to the nationally representative normative sample. There were small but significant positive correlations between age and the scaled scores (ranging from .11 to .18), which raises the possibility that age relations in the current study may be underestimated, because relative to their age peers, the older adults had higher levels of functioning than the younger adults.

Longitudinal participants are often found to have higher scores in various cognitive tests at the initial occasion than individuals who do not return for a second occasion. This pattern was also evident in earlier analyses of VCAP data, with the exception of younger adults who, perhaps because many of the highest functioning individuals were more likely to have moved, exhibited the reverse pattern (Salthouse, 2010a).

Measures

The tests were recently described in Salthouse (2012a) as follows: “Sixteen different cognitive tests selected to represent five cognitive abilities were administered at each occasion (see Salthouse, 2004; Salthouse & Ferrer-Caja, 2003; Salthouse, Pink, & Tucker-Drob, 2008). Vocabulary ability was assessed with Definition Vocabulary, Picture Vocabulary, Synonym Vocabulary, and Antonym Vocabulary tests; reasoning ability was assessed with a Matrix Reasoning test, a Series Completion test, and a test requiring identification of unique letter sets; spatial-visualization ability was assessed with Spatial Relations, Paper Folding, and Form Boards tests; memory ability was assessed with Word Recall, Paired Associates, and Logical Memory tests; and perceptual-speed ability was

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

assessed with Digit Symbol, Pattern Comparison, and Letter Comparison tests. Prior research established that the tests were all reliable, and valid in the sense that they had moderate to high loadings on their respective ability factors; very similar patterns in factor loadings for adults of different ages indicated that the tests had equivalent validity across ages (e.g., Salthouse, 2004; Salthouse & Ferrer-Caja, 2003; Salthouse et al., 2008)."

All participants completed three sessions within a period of about two weeks at each occasion. The sessions were designated by occasion and session, with the first session on the first occasion labeled T11, the second session on the first occasion labeled T12, the first session on the second occasion labeled T21, etc. The test versions performed on successive sessions were identical in format, but involved different items. Because the across-version correlations were greater than .8 for composite scores formed by averaging *z*-scores for the tests representing each ability, and close to 1.0 for latent constructs (Salthouse, 2012b), the different versions can be assumed to be nearly identical in terms of the dimensions of individual differences that were assessed.

The participants with only one version of the relevant tests on the first occasion also completed three sessions within a period of about two weeks. However, different types of tests were performed on the second (T12) and third (T13) sessions. The specific tests varied across test years, but because they all differed from the VCAP reference tests, the participants performing other tests in the T12 and T13 sessions were combined into a single group distinguished from the other group of participants by having performed only one version instead of three versions of the reference tests on the first occasion.

Results[1]

In order to facilitate comparisons across tests, all scores were converted into *z*-score units based on the distribution of T11 scores. Some of the analyses were based on composite scores formed by averaging the *z*-scores of the tests representing the respective cognitive ability. Initial analyses examined linear and quadratic age relations on the T2-T1 difference scores in the groups with one and three T1 versions. With the exception of reasoning ability in the group with one T1 version, all linear age relations were significant, but none of the quadratic age relations was significant.

Two primary sets of analyses were conducted; analyses of variance (ANOVAs) on the longitudinal changes in composite scores with age treated as a categorical variable by grouping participants into age decades, and latent change analyses with age treated as a continuous variable. Because the groups with one and three T1 versions differed in self-rated health, the word recall scaled score, and in the length of the T1-T2 interval (cf. Table 1), these variables were included as covariates in the analyses of variance. (Results without the covariates were nearly identical, with the same pattern of significant and non-significant relations.) A multivariate analysis of variance (MANOVA) with the five composite score differences as dependent variables revealed significant effects of age decade and number of T1 versions, but no significant interaction, and therefore separate analyses were conducted with each ability composite score.

Results of the ANOVAs on the T21-T11 differences in the composite scores are presented in Table 2, where it can be seen that there were significant effects of age decade on all abilities except reasoning. There were also significant effects of number of versions at T1 on the longitudinal change in memory, speed, and spatial visualization abilities, but not in vocabulary or reasoning abilities. Of primary interest were the interactions of age decade

[1]Because of the moderately large sample size, an alpha level of .01 was used in all significance tests.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

and number of versions, and all of these interactions were very small, and not significant for any ability. A post-hoc power analysis based on the smallest sample of 1093 (corresponding to the smallest sample across the five composite scores) revealed that the power to detect a small (f=.10) effect size was .91 with an alpha of .05, and .77 with an alpha of .01.

Similar analyses conducted on the scores from the individual tests revealed that there were significant effects of number of T1 versions on two of the three tests each in with memory, speed, and spatial visualization ability, but that the age by number of T1 versions interaction was not significant in any of the tests.

In order to examine possible interactions with interval, the analyses were repeated with the T1-T2 interval as another factor (coded into 5 discrete categories) instead of as a covariate. The pattern of results with the number of T1 versions factor was very similar to that reported in Table 2. Importantly, there were also no interactions of longitudinal interval with number of T1 versions, or with age and number of T1 versions. At least within the range of T1-T2 intervals examined here, these results provide no evidence that additional experience at the first occasion affects the rate at which change varies as a function of the length of the interval between occasions.

Figure 1 illustrates the longitudinal changes (i.e., T21-T11 differences in composite scores) for participants with one or three T1 versions in each cognitive ability as a function of age decade. Although the figure reveals smaller differences in the composite memory score changes at older ages, the decade X number-of-T1 versions interaction was not significant either for the composite score (Table 2), or for any of the three variables comprising the composite score (i.e., the F's were all less than 1.1, and the partial η^2 values were each less than .004 for the word recall, paired associates and logical memory variables).

Latent change models (Ferrer & McArdle, 2010) were also conducted with the cognitive ability constructs defined by the three or four (for vocabulary) test scores representing each construct, and age, number of T1 versions, and their interactions as predictors of the level and change constructs. Results of these analyses, in the form of the regression coefficients for each predictor, are presented in Table 3. It can be seen that increased age was associated with significantly more negative change for every ability except reasoning. Consistent with the ANOVA results, there were positive effects of the number of T1 versions on the changes in memory, speed, and spatial visualization, but not for the changes in reasoning or vocabulary, and no interactions of age and number of T1 versions on any ability.

Another analysis capitalized on the fact that in 2007 a sample of 227 participants performed the same versions of the tests on the second session of the first occasion and different tests on the third session (Salthouse & Tucker-Drob, 2008). Slightly over half (N = 126) of these individuals returned for a second occasion after an average of 2.3 years, and therefore ANOVAs identical to those reported earlier were conducted with the addition of this third group of participants. The estimated mean changes in the composite scores for each ability in the three groups, adjusted for the covariates, are presented in Table 4. Notice that with the exception of memory and spatial visualization abilities, a single repetition of the same test version had a greater effect on cognitive change than experience with two different versions of the tests. Importantly, as in the other analyses, none of the interactions of decade with amount or type of test experience on the first occasion was significant with the inclusion of this third group.

A final analysis focused on the relation between test experience and longitudinal change at the level of individual differences by examining the relation between the amount of across-session change on the first occasion (i.e., from T11 to T13) and the magnitude of longitudinal change (i.e., from T11 to T21). Specifically, multiple regression analyses were

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

used to predict the T21-T11 difference from age, the T11-T13 difference, and the interaction of age and the T11-T13 difference. The standardized relations for age ranged from −.08 (for vocabulary) to −.32 (for spatial visualization), and all but that for vocabulary was significantly different from zero. All of the relations of within-occasion change with longitudinal change were significant (with a range of standardized coefficients from .35 to .50), but none of the interactions of age and within-occasion change was significant (with a range of standardized coefficients from −.02 to .06).

Discussion

The current study employed a novel approach to investigate the effects of test experience on longitudinal change. That is, instead of trying to estimate the influence of test experience on change with the use of statistical models, or by comparison with a group of test-naïve participants, longitudinal change was examined as a function of different amounts of test experience at the first occasion.

Additional experience with parallel (or identical) versions of cognitive tests on a first occasion affected the magnitude of change on original versions of the test over an interval of approximately 2.5 years for measures of memory, speed, and spatial visualization. Because the across-occasion change was more positive among individuals exhibiting the largest within-occasion improvement, and there was no evidence of differential loss as a function of longitudinal interval, the effects can be inferred to be attributable to a higher level of performance achieved after additional experience with the tasks on the first occasion. That is, the difference from the first occasion to the second occasion was less negative when the final level of performance at the first occasion was higher. An implication of these results is that longitudinal change does not merely reflect factors operating across the interval between occasions, but also, at least indirectly, factors affecting the level of performance on the first occasion.

In all abilities except reasoning in some analyses, between-occasion change was less positive/more negative with increased age. However, none of the interactions of number of T1 versions with age was significant, regardless whether age was a categorical variable in the ANOVAs or a continuous variable in the latent change analyses. Moreover, the failure to find significant interactions is unlikely to be attributable to low power because the analyses had power of at least .77 to detect small effects, the functions in Figure 1 were generally parallel, and there was a consistent pattern in the analyses across measures from individual tests, composite scores, and latent constructs, and at the level of individual differences in the relations between within-occasion (T1) change and between-occasion (T1 to T2) change. The lack of age differences in test experience effects suggests that test experience influences on cognitive change are similar across most of adulthood.

The test experience patterns varied across different cognitive abilities. Little effect of additional first occasion experience was apparent on change in vocabulary, perhaps because of lack of transfer across different vocabulary items, and minimal strategic influences in these tests. There were also small test experience effects with reasoning, which may reflect the difficulty of acquiring strategies in these tests. The large test experience benefits with the relatively unfamiliar spatial visualization tests may reflect the acquisition of appropriate strategies for folding, segmenting, or rotating parts.

As with any project, this study has a number of limitations. One limitation is that most of the participants were healthy adults under 80 years of age at the first occasion. Weaker effects of additional first occasion experience might be evident at older ages, or in people with various types of health problems. The average longitudinal intervals were also relatively

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

short, which may have contributed to larger test experience effects than what might be found in studies with longer intervals. The failure to find age differences in test experience effects on the changes in vocabulary and reasoning may have been attributable to the small main effects of test experience on those variables. Finally, effects of T1 experience were only examined on the second (T2) occasion, and it is not known whether effects of additional experience at the initial occasion might also be detectable on later occasions.

To summarize, results from different types of approaches are converging on a conclusion that practice or retest contributions to change in several cognitive abilities appear to be nearly the same magnitude in healthy adults between about 20 and 80 years of age. These findings imply that age comparisons of longitudinal change are not confounded with differences in the influences of retest and maturational components of change, and that measures of longitudinal change may be underestimates of the maturational component of change at all ages.

Acknowledgments

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.

References

Ferrer E, McArdle JJ. Longitudinal modeling of developmental changes in psychological research. Current Directions in Psychological Science. 2010; 19:149–154.

Ferrer E, Salthouse TA, McArdle JJ, Stewart WF, Schwartz BS. Multivariate modeling of age and retest in longitudinal studies of cognitive abilities. Psychology and Aging. 2005; 20:412–422. [PubMed: 16248701]

Ferrer E, Salthouse TA, Stewart WF, Schwartz BS. Modeling age and retest processes in longitudinal studies of cognitive abilities. Psychology and Aging. 2004; 19:243–259. [PubMed: 15222818]

Folstein MF, Folstein SE, McHugh PR. Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research. 1975; 12:189–198. [PubMed: 1202204]

Kline, RB. Principles and practice of structural equation modeling. 2. New York: Guilford Press; 2005.

McArdle JJ, Ferrer-Caja E, Hamagami F, Woodcock RW. Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. Developmental Psychology. 2002; 38:115–142. [PubMed: 11806695]

Rabaglia C, Salthouse TA. Natural and constrained language production as a function of age and cognitive abilities. Language and Cognitive Processes. 2012; 26:1505–1531.

Salthouse TA. Localizing age-related individual differences in a hierarchical structure. Intelligence. 2004; 32:541–561.

Salthouse TA. When does age-related cognitive decline begin? Neurobiology of Aging. 2009; 30:507–514. [PubMed: 19231028]

Salthouse TA. Influence of age on practice effects in longitudinal neurocognitive change. Neuropsychology. 2010a; 24:563–572. [PubMed: 20804244]

Salthouse TA. Is flanker-based inhibition related to age? Identifying specific influences of individual differences on neurocognitive variables. Brain and Cognition. 2010b; 73:51–61. [PubMed: 20303636]

Salthouse TA. Effects of age on time-dependent cognitive change. Psychological Science. 2011; 22:682–688. [PubMed: 21467547]

Salthouse TA. Does the level at which cognitive change occurs change with age? Psychological Science. 2012a; 23:18–23. [PubMed: 22138158]

Salthouse TA. Does the direction and magnitude of cognitive change depend on initial level of ability? Intelligence. 2012b; 40:352–361. [PubMed: 22711949]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Salthouse TA, Atkinson TM, Berish DE. Executive functioning as a potential mediator of age-related cognitive decline in normal adults. Journal of Experimental Psychology: General. 2003; 132:566–594. [PubMed: 14640849]

Salthouse TA, Berish DE, Siedlecki KL. Construct validity and age sensitivity of prospective memory. Memory & Cognition. 2004; 32:1133–1148. [PubMed: 15813495]

Salthouse TA, Ferrer-Caja E. What needs to be explained to account for age-related effects on multiple cognitive variables? Psychology and Aging. 2003; 18:91–110. [PubMed: 12641315]

Salthouse TA, Pink JE, Tucker-Drob EM. Contextual analysis of fluid intelligence. Intelligence. 2008; 36:464–486. [PubMed: 19137074]

Salthouse TA, Schroeder DH, Ferrer E. Estimating retest effects in longitudinal assessments of cognitive functioning in adults between 18 and 60 years of age. Developmental Psychology. 2004; 40:813–822. [PubMed: 15355168]

Salthouse TA, Siedlecki KL, Krueger LE. An individual differences analysis of memory control. Journal of Memory and Language. 2006; 55:102–125.

Salthouse TA, Tucker-Drob EM. Implications of short-term retest effects for the interpretation of longitudinal change. Neuropsychology. 2008; 22:800–811. [PubMed: 18999354]

Schaie, KW. Internal validity threats in studies of adult cognitive development. In: Howe, ML.; Brainerd, CJ., editors. Cognitive Development in Adulthood: Progress in Cognitive Development Research. N.Y: Springer-Verlag; 1988. p. 241-272.

Siedlecki KL, Salthouse TA, Berish DE. Is there anything special about the aging of source memory? Psychology and Aging. 2005; 20:19–32. [PubMed: 15769211]

Wechsler, D. Wechsler Adult Intelligence Scale. 3. San Antonio, TX: Psychological Corporation; 1997a.

Wechsler, D. Wechsler Memory Scale. 3. San Antonio, TX: Psychological Corporation; 1997b.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

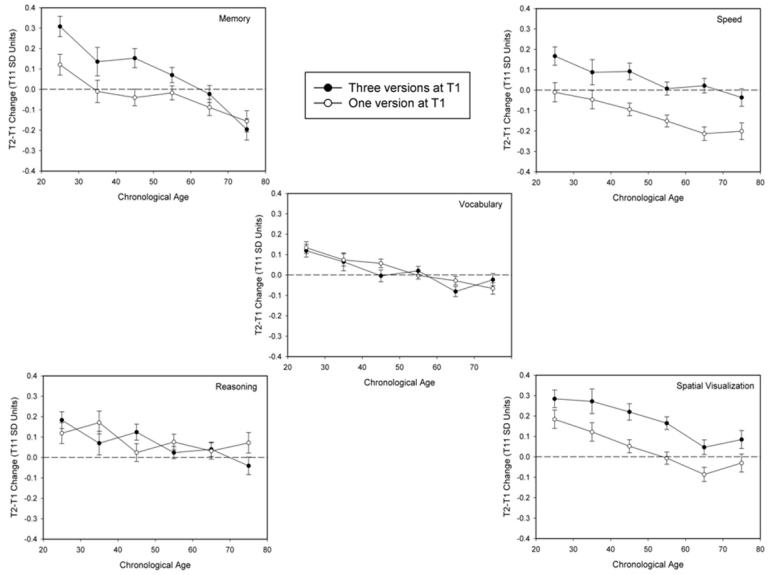

Figure 1.
Means and standard errors of T1 to T2 change in composite scores for each cognitive ability in successive decades for participants performing one or three versions of the cognitive tests at the first occasion.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Sample characteristics

	One T1 version	Three T1 versions	*d*	p
N	878	685		
Age	51.1 (15.1)	52.1 (16.1)	.06	.22
Proportion Female	.67	.67	NA	NA
Self-rated health	2.0 (0.8)	2.3 (0.9)	.36	.00
Years of Education	15.8 (2.6)	15.7 (2.6)	0	.63
T2 MMSE	28.8 (1.5)	28.6 (1.5)	−.13	.11
Scaled Scores				
Vocabulary	12.8 (3.0)	12.7 (2.9)	−.03	.48
Digit Symbol	11.5 (2.9)	11.4 (2.7)	−.04	.33
Logical Memory	12.1 (2.8)	11.9 (2.8)	−.07	.15
Word Recall	12.7 (3.1)	12.2 (3.5)	−.15	.00
T11-T21 (years)	2.9 (1.2)	2.5 (0.9)	−.37	.00

Note: Values in parentheses are standard deviations. NA indicates that the value was not applicable. Health was a self-rating on a scale from 1 for “excellent” to 5 for “poor”. MMSE is the Mini Mental State Exam (Folstein, Folstein & McHugh, 1975). Scaled scores are adjusted for age and have means of 10 and standard deviations of 3 in the nationally representative normative samples (Wechsler, 1997a, 1997b). The final column contains Cohen’s *d* estimates of effect size.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 2

Results of analyses of variance on T21-T11 change in composite scores as a function of age decade and number of test versions at T1

	Decade		Number T1 Versions		Decade X Num. T1 Versions	
	F	$p\eta^2$	F	$p\eta^2$	F	$p\eta^2$
Memory	11.81*	.046	16.24*	.013	1.01	.004
df	5,1233		1,1233		5,1233	
Speed	5.99*	.020	48.86*	.032	0.52	.002
df	5, 1464		1,1464		5, 1464	
Reasoning	2.30	.011	0.39	.000	1.91	.009
df	5,1073		1,1073		5,1073	
Spatial Visualization	11.53*	.042	33.67*	.025	0.39	.001
df	5, 1325		1,1325		5,1325	
Vocabulary	12.15*	.041	0.03	.000	1.51	.005
df	5,1431		1,1431		5,1431	

Note:

* p<.01. The degrees of freedom (df) differed across abilities because not all participants performed every test. Covariates in the analyses were self-rated health, word recall scaled score, and T1-T2 interval. Effect size is represented by partial eta squared ($p\eta^2$) values.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 3

Unstandardized influences on estimates of latent level and latent change from T11 to T21

	Latent Change Estimates			Fit Statistics	
	Age	**# T1 Versions**	**Age X # T1 Versions**	**CFI**	**RMSEA**
Ability					
Memory					
Level	−.010*	−.071*	−.001	.989	.049
Change	−.005*	.061*	−.001		
Speed					
Level	−.025*	−.051*	.000	.987	.060
Change	−.005*	.091*	.001		
Reasoning					
Level	−.018*	.047	.001	.972	.085
Change	−.001	−.014	−.001		
Spatial Visualization					
Level	−.016*	−.011	−.001	.977	.080
Change	−.005*	.082*	.000		
Vocabulary					
Level	.021*	−.031	−.001	.974	.080
Change	−.005*	.005	.001		

Note:

* p<.01. CFI refers to the Confirmatory Fit Index, and RMSEA refers to root mean square error of approximation. CFI values above .90 and RMSEA values below .10 are often considered to represent good fit (Kline, 2005).

Dev Psychol. Author manuscript; available in PMC 2014 November 01.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 4

Estimated mean differences (and standard errors) from T11 to T21 in composite scores with different types of experience at T1, controlling for age, self-rated health, word recall scaled score, and length of T1-T2 interval

	Different Test Versions		Same Test Version
Number of T1 Experiences	**1**	**3**	**2**
Memory	−.051 (.022)	.063 (.022)	.069 (.051)
Speed	−.129 (.019)	.043 (.020)	.143 (.047)
Reasoning	.082 (.021)	.066 (.019)	.107 (.043)
Spatial Visualization	.036 (.017)	.176 (.019)	.137 (.045)
Vocabulary	.016 (.013)	.017 (.014)	.112 (.033)

Note: Data for the same test version are from Salthouse and Tucker-Drob (2008).