Developmental Psychology
2015, Vol. 51, No. 9, 1262–1270

0012-1649/15/$12.00 http://dx.doi.org/10.1037/dev0000030

Test Experience Effects in Longitudinal Comparisons of Adult Cognitive Functioning

Timothy Salthouse
University of Virginia

It is widely recognized that experience with cognitive tests can influence estimates of cognitive change. Prior research has estimated experience effects at the level of groups by comparing the performance of a group of participants tested for the second time with the performance of a different group of participants at the same age tested for the first time. This twice-minus-once-tested method was adapted in the current study to derive estimates of test experience at the level of individual participants. Among the major findings were that experience estimates were smaller at older ages, with measures of vocabulary and speed compared to measures of memory, reasoning, and spatial visualization, and with longer intervals between the first and second occasion. Although relations of overall cognitive ability with test experience effects were weak, there were significant correlations among the experience estimates in different cognitive domains. These results imply that at least in adulthood, simple measures of cognitive change likely underestimate maturational influences on cognitive functioning, and to a greater extent in young adults than in older adults.

Keywords: aging, cognition, longitudinal, practice, retest

A key question in longitudinal research involving adults is how much of the across-occasion change in measures of cognitive functioning is positive, and potentially experience-based, and how much is negative, and likely attributable to factors associated with maturation. Not only is this information critical in accurately determining relations between age and cognitive functioning, but it is also important when evaluating potential correlates of change because individual differences in the effects of test experience could distort observed relations between cognitive change and other variables (e.g., Ferrer, Salthouse, McArdle, Stewart, & Schwartz, 2005).

Two major approaches have been used to investigate influences of test experience on cognitive change; statistical and empirical. Statistical approaches are based on postulating different patterns of influences associated with maturation and with experience, and then attempting to estimate both sets of influences in the same analyses. For example, when there are three or more measurement occasions a continuous linear function could be postulated for maturational influences and a step function could be postulated for experiential influences, with a large gain from the first to the second occasion followed by little or no further increases on subsequent occasions (e.g., Ferrer, Salthouse, Stewart, & Schwartz, 2004, 2005; Granholm, Link, Fish, Kraemer, & Jeste 2010; McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002; Rabbitt, Diggle, Smith, Holland, & McInnes, 2001, Rabbitt, Diggle, Holland, & McInnes 2004). Although frequently used, a limitation of the existing statistical approaches is the absence of independent evidence for the existence of distinct components and their linkage to the hypothesized influences.

A variety of empirical approaches have been used to estimate test experience effects in longitudinal studies. What may be the simplest approach is based on the assumption that all positive change observed in adults is attributable to effects of prior test experience. If this assumption is valid, the gain in test scores from one occasion to the next can serve as an estimate of the effects of experience. However, it is only a lower-bound estimate because the actual experience effect could be larger if, as seems likely, there are also declines across occasions and the observed changes are a mixture of experiential increases and maturational decreases.

Another empirical approach used to investigate experiential effects involves a comparison of change in the same individuals across short and long intervals between assessments. The rationale is that the change over a short interval might be inferred to be completely determined by test experience influences, whereas change over longer intervals would likely also be influenced by maturational factors. However, an intriguing result from a recent comparison of this type (Salthouse, 2013a) was that the correlations between change over short intervals (averaging about 7 days) and change over longer intervals (averaging about 3 years) were negative. Furthermore, increased age was associated with more positive change over short intervals, but with more negative change over longer intervals. These findings imply that different factors may be contributing to the two types of change, and that performance changes over short time scales may not be meaningful as a proxy for experiential effects operating across longer time scales.

This article was published Online First June 22, 2015.

This research was supported by National Institute on Aging Grant R37AG024270. 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.

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

Other empirical approaches have been based on comparisons of longitudinal change in individuals differing in experience-relevant characteristics. For example, one study compared longitudinal change in participants who had completed either one or three versions of the cognitive tests on the initial occasion (Salthouse, 2013b). As one might expect, the longitudinal change was more negative for participants with only one test version at the first occasion than for participants with three versions.

Two additional empirical approaches to estimate effects of prior test experience in longitudinal studies were originally introduced in research on cognitive aging by Schaie and his colleagues. The quasi-longitudinal method (e.g., Salthouse, 2013c, 2014a; in press; Schaie, Labouvie, & Buech 1973; Schaie & Strother, 1968) is based on comparisons of performance of people from the same birth cohort who are tested for the first time in different years, when they were at different ages. The difference in test scores across the same interval as a longitudinal study therefore provides an estimate of cognitive change without prior test experience, and the difference between the observed longitudinal change and the quasi-longitudinal change can serve as an estimate of the test-experience effect over that interval. Comparisons of this type have revealed that the estimates of test experience were positive, and that the quasi-longitudinal age trends were generally more similar to the cross-sectional age trends than to the longitudinal age trends (see Salthouse, 2013c, 2014a, in press).

The twice-minus-once-tested method can be viewed as more direct than the quasi-longitudinal method for the purpose of estimating test experience effects because the difference in scores of a longitudinal sample at the second occasion and the scores of a sample from the same birth cohort tested for the first time, after an adjustment for selectivity of the longitudinal sample, serves as an estimate of the experience effect. Variants of this method have been used in several studies (e.g., Anstey, Sargent-Cox, Garde, Cherbuin, & Butterworth, 2014; Rönnlund & Nilsson, 2006; Ronnlund, Nyberg, Backman, & Nilsson, 2005; Salthouse, 2009, 2010, 2014a), and in almost every case they have yielded positive estimates of the experience effects.

Although a considerable amount of research has focused on the nature of experience effects on cognitive change, many questions remain. The goals of the current study were to use the twice-minus-once-tested method to investigate a number of questions related to test experience effects in longitudinal studies. Unlike prior studies, which were based on group-level comparisons, experience estimates were derived at the level of individual participants, which allowed statistical evaluation of the significance of experience effects.

One question was whether experience effects vary across cognitive domains, with larger effects in certain ability domains than in others. Both Ferrer et al. (2004) and Salthouse (2010) found larger experience estimates for tests of memory tests than for tests of speed, but in neither study was it possible to provide direct comparisons of the effects in different domains. A second question was whether experience effects vary as a function of age in healthy adults, perhaps with smaller effects at older ages (e.g., Salthouse, 2010). A third question capitalized on the availability of a measurement burst design, in which multiple assessments are obtained at each longitudinal occasion, and asked whether across-occasion experience effects were reduced if additional test sessions were administered on each occasion. The fourth question was whether test experience effects were restricted to the interval between the first and second occasion, or whether they are also evident across the interval between the second and third occasions. As noted above, some statistical models assume that experience effects asymptote after the second occasion, but there is apparently no independent evidence validating this assumption.

The preceding questions were investigated with mixed effects analyses of variance (ANOVAs) in which age decade was a between-subjects factor, and either session or interval were within-subjects factors. In addition to the main effects, the interactions involving age decade were of particular interest in indicating whether either the session or interval effects varied as a function of age.

Analyses were also conducted on latent variables created by using the experience estimates from the three sessions at each occasion as indicators of a construct representing the experience effects for that cognitive domain. Four additional questions were investigated with these latent variables. One question was whether the estimates of experience had linear and/or nonlinear relations of age. A second question was whether experience effects were smaller when there were longer intervals between occasions, possibly because information from the first occasion would have had more time to dissipate. A third question was whether test experience effects were larger in individuals with higher levels of cognitive functioning. The final question investigated with experience estimates represented as latent variables was whether people who had large experience effects in one cognitive domain also tended to have large experience effects in other cognitive domains.

A modification of the twice-minus-once-tested procedure described in Salthouse (2010) was used to estimate experience (i.e., practice or retest) effects at the level of individual participants. The procedure is schematically illustrated in Figure 1, and elaborated in Table 1. Solid lines in the figure represent longitudinal change for a given individual, and the dotted lines represent the relation between score on the cognitive variable and age for people tested for the first time in different test years.

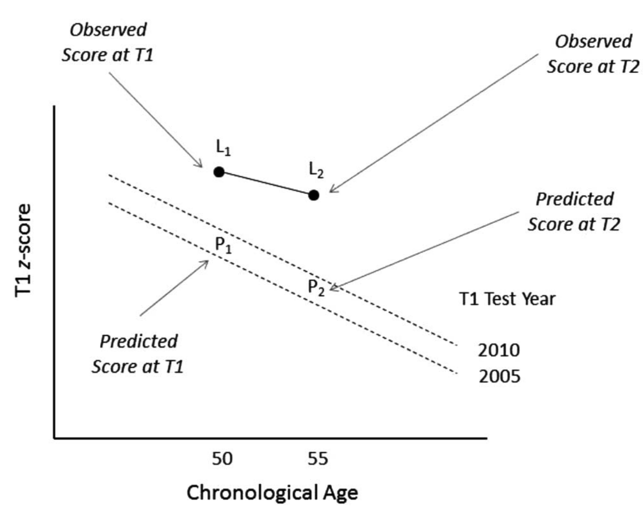

Figure 1. Schematic illustration of method used to estimate experience effects for a given individual. The two ages and two test years are included to make the example more concrete, but the analyses were conducted across multiple ages and multiple test years. See text for details.

Table 1
Terminology and Equations Used to Derive Estimates of Test Experience at the Level of Individuals

P1:	Predicted value from the total sample at the longitudinal T1 age
P2:	Predicted value for participants tested only once at the longitudinal T2 age
L1:	Observed score at T1 for the target longitudinal individual
L2:	Observed score at T2 for the target longitudinal individual
D_{12}:	L2 – P2 (twice-minus-once-tested difference)
S_{12}:	L1 – P1 (selectivity of longitudinal individual)
E_{12}:	D_{12} – S_{12} (experience adjusted for selectivity)

Because of more inclusive recruitment procedures in later test years, the relations between the cognitive composite scores and test year were negative. Influences of test year were therefore partialed from all T1 (Time 1) scores to effectively allow comparisons at the average T1 test year. A regression equation relating the cognitive composite residual score to age was then computed, and regression parameters based on the data from all participants were used to predict the score expected for a participant of the target age on the first longitudinal occasion (designated P1). The regression analyses relating cognitive residual score to age were then repeated with participants who only had one assessment, and the regression parameters from these once-tested participants were used to predict the score of a participant at the target age on the second longitudinal occasion (designated P2).

The rationale for computing experience estimates from these values is summarized in Table 1. D_{12} is the contrast of the individual's score at T2 (Time 2) with the predicted score for once-tested individuals of that age, and it corresponds to the difference on the second occasion for the score of a specific individual and the score for the average individual of that age who was only tested once. S_{12} is the selectivity of the individual's score at T1 relative to the predicted score for all individuals of that age. Finally, E_{12} is the estimated experience effect for the target individual over the T1–T2 interval.

Parallel analyses were conducted to investigate experience effects across the T2–T3 interval, in this case partialing the effects of T2 test year from the T2 cognitive composite scores. Only participants with data for all three occasions were included in these analyses to allow direct comparisons of effects across the T1–T2 and T2–T3 intervals.

Method

Participants

Although participants in the Virginia Cognitive Aging Project ranged up to 99 years of age, the analyses in this report were restricted to adults between 18 and 80 years of age on the first occasion to maximize sample sizes in each age group, and minimize influences of late-life diseases that could affect cognition. The participants were recruited for their initial occasion between 2001 and 2014, with subsequent occasions occurring between 2004 and 2014. The average interval between the first and second occasion was 3.0 years, with a range from 0.2 to 13.1 years. On each occasion the participants reported to the laboratory for three sessions within a period of about 2 weeks, and performed parallel versions of the tests on each session. However, participants in early years of the project performed different tests on the second and third sessions (Salthouse, 2013b), and therefore the sample sizes on Sessions 2 and 3 were smaller than that on Session 1. The procedure used to equate the mean performance across versions was described in Salthouse (2012).

Characteristics of the participants as a function of age decade are reported in Table 2. It can be seen that there was a small relation of age to self-rated health, in the direction of poorer ratings at older ages, but that increased age was associated with more years of education and slightly higher estimated IQ.

Measures

The cognitive tests were described in a recent publication (Salthouse, 2014b, pp. 568–569) as follows:

Table 2
Characteristics of Participants As A Function of Decade

	Decade						
Characteristic	20s	30s	40s	50s	60s	70s	Age correlation
Number of occasions							
N with 1 +	883	501	779	1144	883	560	NA
N with 2 +	276	223	436	660	512	324	NA
N with 3 +	110	102	259	379	301	165	NA
Age	23.1 (3.2)	34.3 (2.8)	45.0 (2.9)	54.4 (2.8)	64.1 (2.8)	74.2 (2.9)	NA
Prop. Female	.58	.69	.71	.70	.66	.58	.02
Health	2.0 (0.9)	2.1 (0.8)	2.2 (0.9)	2.2 (0.9)	2.1 (0.9)	2.4 (0.9)	.10*
Education	14.7 (2.1)	15.5 (5.9)	15.3 (2.7)	15.6 (4.3)	16.4 (2.8)	15.7 (5.6)	.12*
Est. IQ	109.1 (12.5)	107.6 (15.0)	108.5 (15.4)	109.8 (15.1)	112.0 (13.3)	108.9 (13.7)	.05*
PC1	0.65 (1.0)	0.34 (1.1)	0.23 (1.0)	0.07 (0.9)	−0.11 (0.8)	−0.63 (0.8)	−.36*
T1–T2 Int.	3.1 (2.0)	3.1 (2.0)	3.3 (1.9)	3.0 (1.6)	2.8 (1.5)	2.7 (1.3)	−.09*
T2–T3 Int.	3.1 (1.5)	3.4 (1.6)	3.3 (1.5)	3.1 (1.4)	2.9 (1.2)	3.1 (1.4)	−.07

Note. Health is a self-rating on a scale from 1 (*excellent*) to 5 (*poor*). Est. IQ is an estimate of IQ based on age-adjusted scores on three tests found to be highly related to Wechsler IV full scale IQ (Salthouse, 2014b). PC1 is the first principal component based on scores of the 16 tests in the first session of the first occasion. T1–T2 and T2–T3 intervals are in years. Prop. Female = proportion of females.
* $p < .01$.

A total of 16 cognitive tests, representing five cognitive abilities, were administered in the same order to all participants. Vocabulary was assessed by a provide-the-definition test, a picture naming test, and multiple-choice synonym and antonym tests. Reasoning was assessed by a matrix reasoning test, a letter sets test, and a series completion test. Spatial visualization ability was assessed by a spatial relations test, a paper folding test, and a form boards test. Episodic memory was assessed by word recall, paired associates, and story (logical) memory tests. Perceptual speed ability was assessed by a digit symbol substitution test, pattern comparison test, and letter comparison test.

Cognitive functioning was examined with composite scores formed by averaging *z* scores (based on the means and standard deviations from the first assessment in the complete sample) for the three or four measures representing each cognitive ability. A principal components analysis was conducted on the five composite scores on the first session of the first (T1) and second (T2) occasions. The first principal component (PC1) was associated with 56.9% of the variance, and therefore it was used as a general measure of cognitive functioning in some analyses.

Results

The methods outlined in Figure 1 and Table 1 were used to derive test experience estimates for each individual. The experience estimates (with standard errors) for the composite scores in each ability domain across the T1–T2 interval in the three sessions are plotted as a function of age decade in Figure 2.

The data summarized in Figure 2 were initially analyzed with an ANOVA in which cognitive domain, age decade, and session were factors. The primary interest in the initial analysis were the domain effects, which were significant ($p < .01$) for the main effect of domain ($F = 9.49$, partial eta$^2 = .008$), and for the interactions of Domain $\times$ Decade ($F = 2.70$, partial eta$^2 = .011$), and Domain $\times$ Session ($F = 4.25$, partial eta$^2 = .004$), but the triple interaction was not significant ($F < 1.4$, $p > .06$). Post hoc comparisons revealed no significant differences among the experience estimates for memory, reasoning, and spatial visualization abilities, but each of those was significantly larger than the estimates for speed and vocabulary which were not significantly different from one another.

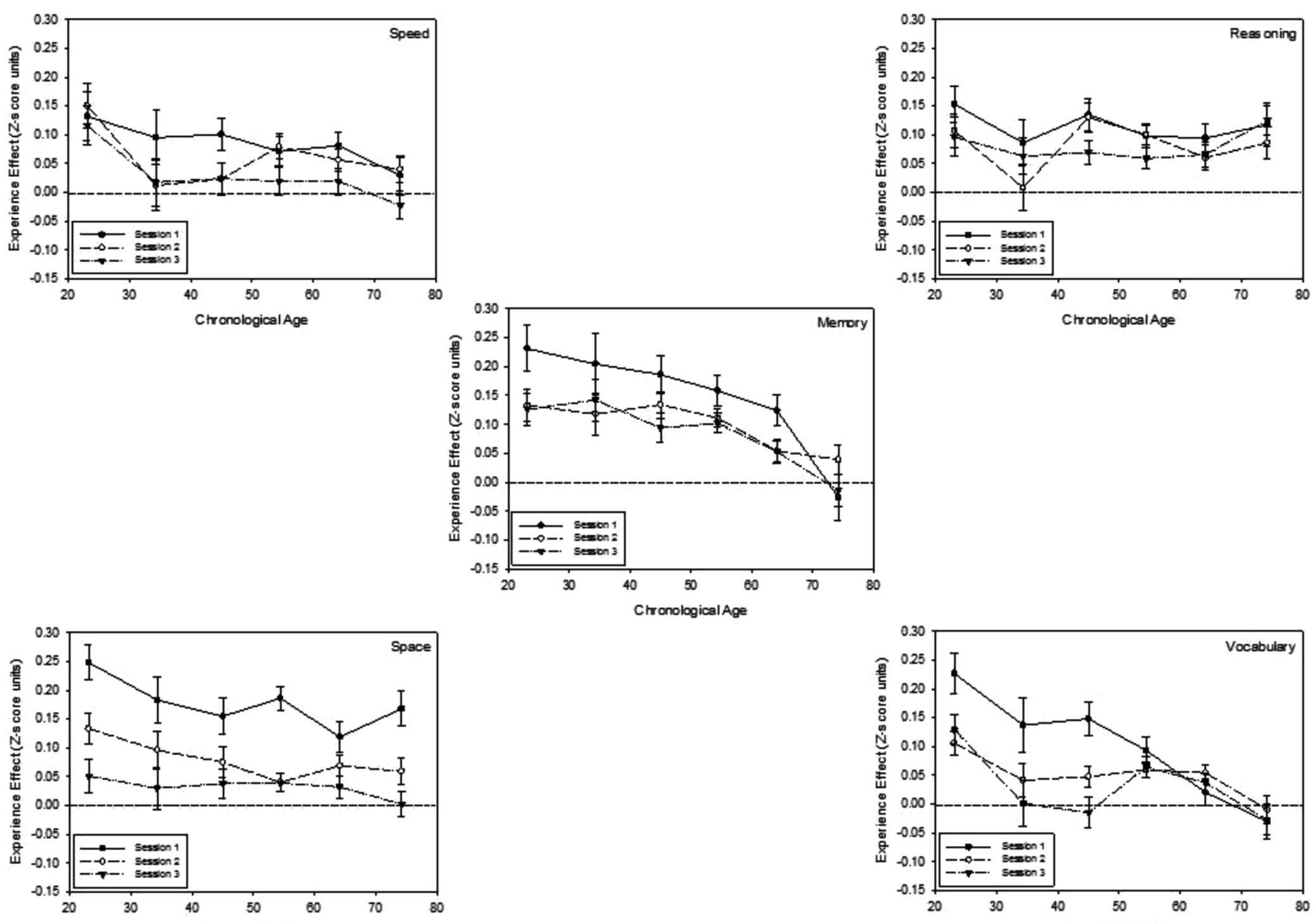

Figure 2. Mean (and standard error) of T1–T2 experience estimates for the composite scores in five cognitive domains in three sessions in different decades.

Results of separate Decade × Session ANOVAs in each cognitive domain are summarized in Table 3. It can be seen that the largest experience estimates were evident on the first session, but with the exception of speed on the third session, the estimates in all sessions were significantly greater than zero. The interaction of age and session was significant only with the vocabulary measures, and Figure 2 indicates that it is attributable to larger experience estimates on the first session at younger ages.

Experience effects across the T1–T2 and T2–T3 intervals were investigated for participants who had completed at least three occasions. Means (and standard errors) of the experience estimates for the T1–T2 and T2–T3 intervals are plotted as a function of age in Figure 3. Only data from the first session in each occasion were included in these analyses because this session had the greatest amount of data, and only one interaction involving session was significant in the analyses of data for the T1–T2 interval. The initial ANOVA with cognitive domain, decade, and interval (i.e., T1–T2 or T2–T3) as factors revealed significant ($p < .01$) effects of domain ($F = 33.39$, partial eta^2 = .025), domain X decade ($F = 2.76$, partial eta^2 = .010), and domain X interval ($F = 6.77$, partial eta^2 = .005), but not the triple interaction ($F < 1.2$, $p > .26$). Post hoc analyses revealed that the experience estimates were largest for spatial visualization and reasoning which did not differ from one another, intermediate for memory, which was significantly larger than speed, which was significantly larger than vocabulary.

Results of separate Decade × Interval ANOVAs in each cognitive domain are summarized in Table 4. Although the experience effects were generally larger across the T1–T2 interval than the T2–T3 interval, the interval difference was significant only for the reasoning and vocabulary measures. Moreover, all of the experience estimates except that for vocabulary across the T2–T3 interval were significantly greater than zero.

Latent Variable Analyses

Although the availability of experience estimates at the level of individuals allows relations of experience with other variables to be examined, the experience estimates are likely to have considerable measurement error, which can attenuate relations with other variables. Additional relational analyses were therefore conducted on latent variables created by using the experience estimates in the three sessions as indicators of experience effects in the relevant cognitive domain. For example, a latent variable representing test experience for memory was created by postulating that it contributed to the experience estimates on Sessions 1, 2, and 3. The latent variable models for each domain had excellent fits to the data, as the fit statistics when age and age^2 were used as predictors of the latent variables were all in the good to excellent range (i.e., χ^2/df <2.5, comparative fit index > .96, and root mean square error of appoximation < .02).

Unstandardized and standardized coefficients from these analyses are reported in Table 5. It can be seen that the experience estimates were less positive at older ages, but that the quadratic age trend was significant only for memory. The pattern in Figure 2 suggests that the nonlinear trend with memory is attributable to an accelerated decrease in the experience effects after the decade of the 60s.

As one might expect, all of the experience estimates were smaller with longer intervals between the first and second longitudinal occasions. However, it is noteworthy that the relations with the T1–T2 interval were much larger than the relations with age. For example, with the memory variable the experience estimate decreased .003 standard deviations per year of age, but decreased .045 standard deviations per year of interval. Similar patterns were evident in the other cognitive domains, although the significant

Table 3
Results of Decade × Session (1, 2, or 3) Analyses of Variance on Experience Estimates for the T1–T2 Interval

		F ratio (partial eta^2)		
	M (*SD*)	Decade	Session	Decade × Session
Memory		9.16* (.037)	9.99* (.008)	1.56 (.007)
Session 1	.146* (.488)			
Session 2	.095* (.315)			
Session 3	.079* (.333)			
Speed		3.00 (.012)	7.23* (.006)	0.89 (.004)
Session 1	.084* (.435)			
Session 2	.062* (.377)			
Session 3	.027 (.383)			
Reasoning		1.22 (.005)	3.75 (.003)	0.87 (.004)
Session 1	.115* (.408)			
Session 2	.086* (.342)			
Session 3	.078* (.332)			
Spatial visualization		1.76 (.007)	50.43* (.041)	1.11 (.005)
Session 1	.172* (.411)			
Session 2	.070* (.311)			
Session 3	.032* (.317)			
Vocabulary		12.19* (.049)	9.38* (.008)	3.12* (.013)
Session 1	.083* (.417)			
Session 2	.052* (.257)			
Session 3	.037* (.330)			

* $p < .01$. The asterisk (*) in the mean column indicates that the value was significantly ($p < .01$) greater than zero.

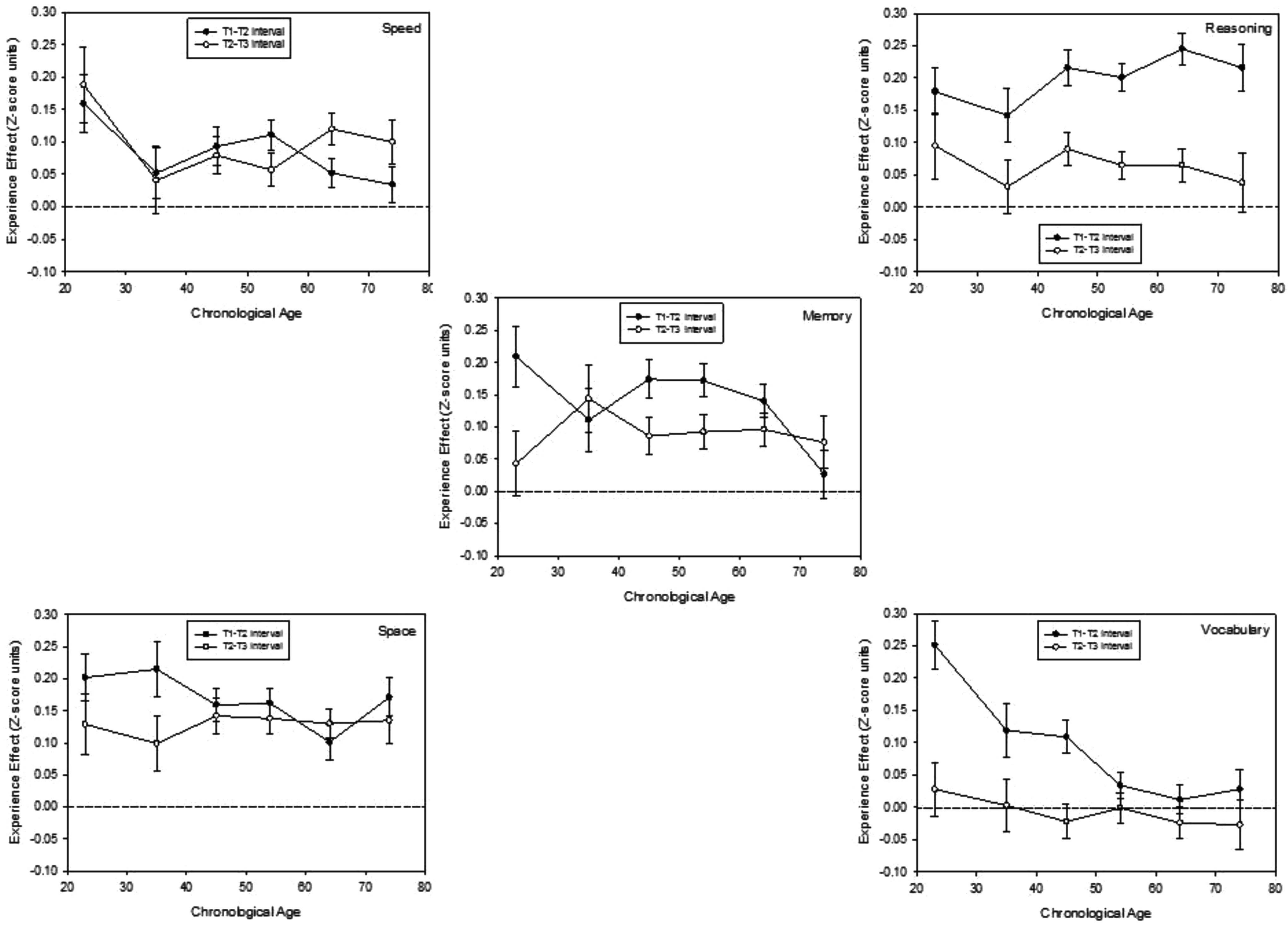

Figure 3. Mean (and standard error) of T1–-T2 and T2–T3 experience estimates for the composite scores in five cognitive domains on the first session in different decades.

Age × Interval interaction in the reasoning domain indicated that with the reasoning tests the effects of interval on test experience were larger at older ages.

Higher levels of general cognitive ability, as assessed both with estimated IQ and with the first principal component, were associated with positive experience estimates in the memory domain and negative experience estimates in the vocabulary domain. These results indicate that compared to lower-ability individuals, higher-ability individuals had somewhat larger experience effects in the memory tests, but smaller experience-related gains in the vocabulary tests. None of the interactions of cognitive ability and age was significant.

To determine whether individuals with large test experience effects in one cognitive domain also had large experience effects in other cognitive domains, correlations were computed between the latent variable experience estimates in each domain. Because the experience estimates were all significantly related to age, influences of age were statistically controlled when computing the correlations. The resulting correlations are summarized in Table 6. Notice that all but one of the correlations was significantly greater than zero, and that the significant correlations ranged from .26 to .66.

Discussion

Statistical methods used to investigate experience effects in earlier studies have been based on the assumption that there are different functions for maturational and experiential components, but the validity of that assumption has not been independently verified. Moreover, even if there were two distinct functions, they may not exclusively represent influences of maturation and experience, and to the same extent in people of different ages. The twice-minus-once-tested method can be considered a more direct method of estimating experience effects, but in the past it has only been applied at the group level, which has precluded statistical comparison of the effects. A novel feature of the current study was the derivation of estimates of test experience at the level of individual participants, which allowed investigation of experience effects across ability domains, successive longitudinal occasions, sessions within occasions, intervals between occasions, and age of the participants.

The results of this study are relevant to each of the questions posed in the introduction. With respect to possible differences across cognitive domains, the experience estimates were larger for the domains of memory, reasoning, and spatial visualization than for speed and vocabulary. These differences may be attributable to

Table 4
Results of Decade × Interval (T1–T2 or T2–T3) Analyses of Variance on Session 1 Experience Estimates

		F ratio (partial eta^2)		
	M (*SD*)	Decade	Interval	Decade × Interval
Memory		2.11 (.008)	3.97 (.003)	1.43 (.005)
T1–T2	.145* (.488)			
T2–T3	.089* (.489)			
Speed		3.33* (.013)	0.40 (.000)	1.19 (.005)
T1–T2	.084* (.428)			
T2–T3	.091* (.477)			
Reasoning		1.38 (.005)	36.03* (.027)	0.47 (.002)
T1–T2	.210* (.426)			
T2–T3	.066* (.456)			
Spatial visualization		1.24 (.005)	2.97 (.002)	0.73 (.003)
T1–T2	.156* (.433)			
T2–T3	.133* (.446)			
Vocabulary		9.41* (.035)	19.46* (.015)	1.59 (.006)
T1–T2	.068* (.408)			
T2–T3	−.013 (.440)			

* $p < .01$. The asterisk (*) in the mean column indicates that the value was significantly ($p < .01$) greater than zero.

the greater involvement of task-specific strategies with memory, reasoning, and spatial visualization because the tests are relatively unfamiliar compared to tests of speed and vocabulary, and thus may be more amenable to the acquisition of strategies that contribute to higher performance on the second occasion.

The experience estimates were generally smaller at older ages, particularly within the memory domain. The latent variable analyses in which age was a continuous variable were probably more sensitive than the ANOVAs in which age was a categorical variable, and only the estimates with reasoning were not significantly related to age in the latent variable analyses. These findings suggest that increased age is associated with diminished ability to benefit from the prior test experience, which implies that the distortion of longitudinal cognitive change by positive effects of prior experience with the tests will likely be greater among younger adults than among older adults.

The significant effects of session, and of the first versus second longitudinal interval, indicate that amount of experience with the tests is an important factor affecting the magnitude of experience effects. An advantage of the measurement burst design implemented in the current study is that short-term and longer-term experience effects can be distinguished in the within-occasion and between-occasion contrasts. In conventional longitudinal studies there is only a single measurement at each occasion, and thus only effects over long intervals can be investigated.

The across-session within-occasion comparisons revealed smaller experience estimates across the T1–T2 interval on Sessions 2 and 3 than on Session 1. However, it is noteworthy that the experience estimates were greater than zero on the second and third sessions, and therefore additional experience within each occasion did not completely eliminate the positive experience effects. The across-occasion experience comparisons revealed that the experience estimates were significantly smaller across the T2–T3 interval than across the T1–T2 interval for memory and vocabulary. However, even within these domains the experience estimates were not zero in the T2–T3 interval, and thus it is not the case that test experience effects are restricted to the interval between the first and second occasion.

Table 5
Unstandardized (and Standardized) Relations Between T1–T2 Latent Experience Variables

Variable	Memory	Speed	Reasoning	Spatial visualization	Vocabulary
Age	−.003* (−.260)	−.001* (−.111)	.000 (.033)	−.001* (−.119)	−.004* (−.403)
Age^2	.000* (−.141)	.000 (.035)	.000 (.032)	.000 (.076)	.000 (.009)
T1–T2 Interval	−.045* (−.328)	−.054* (−.379)	−.028* (−.277)	−.036* (−.258)	−.019* (−.185)
Age × Interval interaction	.000 (−.018)	.000 (−.042)	−.001* (−.126)	.000 (−.019)	−.001 (−.089)
Est. IQ	.002* (.101)	.000 (−.031)	−.001 (−.057)	.000 (.026)	−.011* (−.256)
Age × Est. IQ interaction	.000 (−.024)	.000 (−.054)	.000 (.043)	.000 (−.049)	.000 (.020)
PC1	.039* (.169)	.006 (.024)	.008 (.040)	.017 (.116)	−.133* (−.221)
Age × PC1 interaction	.001 (.068)	−.001 (−.067)	.001 (.052)	.000 (−.045)	.001 (.017)

Note. Est. IQ = estimate of IQ based on age-adjusted scores on three tests found to be highly related to Wechsler IV full scale IQ (Salthouse, 2014b); PC1 = first principal component.
* $p < .01$.

Table 6
Correlations Among Age-Partialled T1–T2 Latent Experience Variables Across Different Cognitive Domains

Variable	1	2	3	4	5
1. Memory	—	.33*	.48*	.35*	.66*
2. Speed		—	.44*	.35*	.26*
3. Reasoning			—	.62*	.54*
4. Spatial visualization				—	.16
5. Vocabulary					—

Note. Model fit: $\chi^2/df = 1.367$; comparative fit index = .973; root mean square error of approximation = .009.
* $p < .01$.

Because there was variability across participants in the length of the interval between the first and second occasions, it was possible to investigate the effect of the interval between occasions on the test experience estimates. The results in Table 4 indicate that the magnitude of the experience effect decreased with increases in the length of the T1-T2 interval. It is not surprising that the benefits associated with prior exposure to the tests would dissipate over time, but it is noteworthy that there was no evidence of age differences in the effects of interval on the experience estimates. These results suggest that the processes involved in the existence of the test experience effects are somewhat independent of the processes contributing to the decay of those effects over time because age was associated with the former but not the latter.

The question of the relation of overall cognitive ability to test experience effects is interesting because test experience effects might be assumed to be related to learning, which would be expected to be more efficient in individuals with higher levels of cognitive ability. The cognitive ability relations were positive for the memory tests, possibly because higher-ability individuals were more effective at forming relevant strategies. However, the test experience effects with the vocabulary tests were smaller for higher-ability individuals than lower-ability individuals, which may be attributable to greater initial familiarity of the higher-ability individuals with accessing and retrieving word meanings, thereby resulting in a smaller benefit from the first test experience than the lower-ability individuals.

Finally, the discovery of significant correlations of the test experience effects across cognitive domains suggests that there is a common influence on the test experience effects in different domains. The results just described indicate that the common influence does not appear to be related to general cognitive ability, and instead it may be associated with factors such as greater test sophistication or reduced test anxiety.

It is worth considering implications of the current results for how one might deal with experience effects in longitudinal studies of cognitive functioning. One possible approach is to restrict comparisons to cognitive domains exhibiting the smallest experience effects, such as vocabulary and speed. However, this approach is not desirable because it would result in ignoring important cognitive domains, and it would also fail to eliminate all experience effects as positive experience effects were still evident with vocabulary and speed. A second less-than-optimal approach is to provide additional experience on the first occasion to try to eliminate the experience effects. The discovery of positive experience effects in the longitudinal contrasts on the second and third sessions suggests that this approach is unlikely to be successful because the across-occasion experience effects appear to be somewhat independent of within-occasion experience effects. Furthermore, ignoring data on the first occasion and focusing only on change across subsequent occasions would not be an effective method of eliminating experience-related influences on change because significant experience effects were still apparent across the T2–T3 interval.

A more promising possibility might consist of increasing the length of the interval between measurement occasions to allow test experience effects to dissipate. Indeed, it is possible to compute the number of years needed for the experience effect to reach a value of zero based on the average experience effect, and the slope relating the experience effect to the interval between occasions. As an example, in the memory domain the average experience effect across the T1–T2 interval of 3 years was .146 standard deviations. Dividing this value by the slope of −.045 standard deviations per year of interval results in an estimate of about 6.2 years (i.e., 3 to reach the .146 level plus [.146/.045]) for the average experience effect to decrease to zero. Although theoretically plausible, a limitation of this approach is that it would not be useful if one is interested in change over shorter intervals.

Finally, statistical models of experience effects may be feasible, but the results of this study suggest that the models will have to be more complex than those used in the past. Not only is it unrealistic to assume a simple step function for experience with effects only across the first interval, but it is also important to consider the cognitive domain, the age of the participant, and the length of the interval between occasions.

This study focused on cognitive change in adults, but similar methods of distinguishing experiential and maturational components of change might be productively applied in childhood. That is, prior experience with the tests may not only lead to underestimates of cognitive declines in adulthood, but also to overestimates of cognitive gains in childhood, and the twice-minus-once-tested procedure could be used to investigate this latter possibility.

To conclude, although there is a consensus that cognitive change is a mixture of negative influences associated with maturation or development and positive influences associated with prior experience, the complex nature of the experience effects is not always recognized. However, the results of this study indicate that experiential determinants of change must be accurately assessed to evaluate the contribution of different components to cognitive change.

References

Anstey, K. J., Sargent-Cox, K., Garde, E., Cherbuin, N., & Butterworth, P. (2014). Cognitive development over 8 years in midlife and its association with cardiovascular risk factors. *Neuropsychology, 28,* 653–665. http://dx.doi.org/10.1037/neu0000044

Ferrer, E., Salthouse, T. A., McArdle, J. J., Stewart, W. F., & Schwartz, B. S. (2005). Multivariate modeling of age and retest in longitudinal studies of cognitive abilities. *Psychology and Aging, 20,* 412–422. http://dx.doi.org/10.1037/0882-7974.20.3.412

Ferrer, E., Salthouse, T. A., Stewart, W. F., & Schwartz, B. S. (2004). Modeling age and retest processes in longitudinal studies of cognitive abilities. *Psychology and Aging, 19,* 243–259. http://dx.doi.org/10.1037/0882-7974.19.2.243

Granholm, E., Link, P., Fish, S., Kraemer, H., & Jeste, D. (2010). Age-related practice effects across longitudinal neuropsychological assessments in older people with schizophrenia. *Neuropsychology, 24,* 616–624. http://dx.doi.org/10.1037/a0019560

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology, 38,* 115–142. http://dx.doi.org/10.1037/0012-1649.38.1.115

Rabbitt, P., Diggle, P., Holland, F., & McInnes, L. (2004). Practice and drop-out effects during a 17-year longitudinal study of cognitive aging. *The Journals of Gerontology Series B, Psychological Sciences and Social Sciences, 59,* 84–97. http://dx.doi.org/10.1093/geronb/59.2.P84

Rabbitt, P., Diggle, P., Smith, D., Holland, F., & McInnes, L. (2001). Identifying and separating the effects of practice and of cognitive ageing during a large longitudinal study of elderly community residents. *Neuropsychologia, 39,* 532–543. http://dx.doi.org/10.1016/S0028-3932(00)00099-3

Rönnlund, M., & Nilsson, L.-G. (2006). Adult life-span patterns in WAIS–R Block Design performance: Cross-sectional versus longitudinal age gradients and relations to demographic factors. *Intelligence, 34,* 63–78. http://dx.doi.org/10.1016/j.intell.2005.06.004

Ronnlund, M., Nyberg, L., Backman, L., & Nilsson, L.-G. (2005). Stability, growth, and decline in adult life span development of declarative memory: Cross-sectional and longitudinal data from a population-based study. *Psychology and Aging, 20,* 3–18.

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging, 30,* 507–514. http://dx.doi.org/10.1016/j.neurobiolaging.2008.09.023

Salthouse, T. A. (2010). Influence of age on practice effects in longitudinal neurocognitive change. *Neuropsychology, 24,* 563–572. http://dx.doi.org/10.1037/a0019026

Salthouse, T. A. (2012). Robust cognitive change. *Journal of the International Neuropsychological Society, 18,* 749–756. http://dx.doi.org/10.1017/S1355617712000380

Salthouse, T. A. (2013a). Effects of age and ability on components of cognitive change. *Intelligence, 41,* 501–511. http://dx.doi.org/10.1016/j.intell.2013.07.005

Salthouse, T. A. (2013b). Effects of first occasion test experience on longitudinal cognitive change. *Developmental Psychology, 49,* 2172–2178. http://dx.doi.org/10.1037/a0032019

Salthouse, T. A. (2013c). Within-cohort age-related differences in cognitive functioning. *Psychological Science, 24,* 123–130. http://dx.doi.org/10.1177/0956797612450893

Salthouse, T. A. (2014a). Why are there different age relations in cross-sectional and longitudinal comparisons of cognitive functioning? *Current Directions in Psychological Science, 23,* 252–256. http://dx.doi.org/10.1177/0963721414535212

Salthouse, T. A. (2014b). Selectivity of attrition in longitudinal studies of cognitive functioning. *Journals of Gerontology: Series B*. Psychological Science, *69,* 567–574.

Salthouse, T. A. (in press). Aging cognition unconfounded by prior test experience. *Journals of Gerontology, Series B. Psychological Science*.

Schaie, K. W., Labouvie, G. V., & Buech, B. U. (1973). Generational and cohort-specific differences in adult cognitive functioning: A fourteen-year study of independent samples. *Developmental Psychology, 9,* 151–166. http://dx.doi.org/10.1037/h0035093

Schaie, K. W., & Strother, C. R. (1968). The effect of time and cohort differences on the interpretation of age changes in cognitive behavior. *Multivariate Behavioral Research, 3,* 259–293. http://dx.doi.org/10.1207/s15327906mbr0303_1

Received November 18, 2014
Revision received March 30, 2015
Accepted May 20, 2015 ■