Intelligence 48 (2015) 51–57

Contents lists available at ScienceDirect

Intelligence

Implications of the Flynn effect for age–cognition relations

Timothy A. Salthouse

Department of Psychology, University of Virginia, Charlottesville, VA 22904, USA

ARTICLE INFO

Article history:
Received 28 August 2014
Received in revised form 3 October 2014
Accepted 16 October 2014
Available online xxxx

Keywords:
Flynn effect
Longitudinal
Cross-sectional
Aging
Cognition

ABSTRACT

Many studies have documented that cognitive performance is often higher among people of the same age who are tested in more recent years, and it is sometimes suggested that this phenomenon will distort the relations between age and cognition in cross-sectional studies. This possibility was examined with data from two large projects involving adults across a wide age range. The results indicated that there were similar time-of-measurement increases in cognitive scores at different ages, which were accompanied by nearly constant cross-sectional age differences, but positively inflated estimates of longitudinal age differences. It is proposed that when the Flynn effect is of comparable magnitude in adults of different ages, longitudinal comparisons of age–cognition relations are more subject to distortion than cross-sectional comparisons.

The Flynn effect (named after James Flynn but originally described many years earlier, cf. Lynn, 2013) refers to the phenomenon that people of the same age who are tested in more recent years tend to have higher scores on cognitive tests than people tested in earlier years. Although many questions remain about the mechanisms for the effect, and its generality across ability domains, cultures, and historical periods, the basic phenomenon has been widely replicated and can be considered to be well established (Trahan, Stuebing, Fletcher, & Hiscock, 2014; Williams, 2013).

A number of researchers have postulated that the Flynn effect will lead to a distortion of cross-sectional relations between age and cognition (e.g., Baxendale, 2010; Hiscock, 2007; Ronnlund & Nilsson, 2009). For example, Flynn stated that "cross-sectional data, as a measure of the effects of aging on IQ, are suspect. … Cross-sectional data compare, for example, 80-year-old subjects with a group of 20-year-old subjects, with both groups being tested at the same time. This makes sense only if current 20-year-olds have the same IQ as 20-year-olds did two generations ago, that is, when today's 80-year-olds were 20" (Flynn, 1987 p. 187).

However, the thesis of this article is that the implications of the Flynn effect for both cross-sectional and longitudinal relations between age and cognition depend on whether the Flynn effect represents a cohort effect or a period (time-of-measurement) effect. Consider the definitions of these terms provided by Schaie (2013)

> "… cohort effects represent the impact of historical effects on a group of individuals who share similar environmental circumstances at equivalent points in their maturation sequence … On the other hand, time-of-measurement effects represent those events that have an impact on all members of the population experiencing a common historical exposure, regardless of cohort membership (p. 25)."

> "Period effects would be ones caused by one or more social innovations that act equally at a point in time on all individuals, regardless of age … A cohort effect would be one acting on children or adults of a particular age, persisting across time (p. 340)."

Based on these definitions, it can be inferred that the distinguishing feature of a period effect is that time-of-measurement influences are similar in people of all ages, and do not vary according to birth year or cohort. Fig. 1 illustrates a situation of this type with cognitive test scores plotted as a function of time

E-mail address: Salthouse@virginia.edu.

http://dx.doi.org/10.1016/j.intell.2014.10.007
0160-2896

of measurement. The thick dashed line represents period influences, which are portrayed as progressively more positive in successive test years. The vertical rectangles correspond to cross-sectional comparisons, and it can be seen that if the period effects are similar at each age, cross-sectional age differences would be expected to be approximately parallel at each test year. That is, even if the absolute level of performance is higher in successive test years, differences in cognitive tests scores between 25-year-old, 45-year-old, and 65-year-old participants would be expected to be comparable in the 2000, 2005, and 2010 test years. In contrast, if the period effects varied with age, possibly with greater time-related improvements at younger ages, cross-sectional age differences would be expected to be larger in more recent test years.

Inspection of Fig. 1 reveals another implication of the Flynn effect for age–cognition relations, namely, that positive time-of-measurement effects can lead to a distortion of longitudinal comparisons. That is, if the factors contributing to higher scores on more recent test years in different people (portrayed by the thick dashed lines) also operate within the same people (portrayed by the dotted diagonal boxes), then longitudinal comparisons will likely be inflated by the presence of the Flynn effect. In other words, because in longitudinal designs assessments at successive ages necessarily occur in more recent years, some of the age-related longitudinal differences in cognitive performance may be attributable to positive period effects. Moreover, although the distortion of the longitudinal comparisons will be larger in later birth cohorts if the period effects are greater at younger ages, some distortion will be evident whenever period effects are positive.

Schaie (e.g., 2013, pp. 192–193) recognized that longitudinal comparisons might be influenced by positive time-of-measurement effects and proposed that adjusted longitudinal change could be estimated by subtracting the estimated time-of-measurement effect from the observed change. Initial analyses of this type were reported in Salthouse (1991), but they were limited by the data available at that time.

There were three major goals of the current study. The first was to investigate whether the magnitude of the Flynn effect was similar at different ages in adulthood. The second goal was to examine cross-sectional comparisons in different test years to determine whether the relative age differences were remaining constant, or increasing over time. The third goal was to estimate longitudinal change after adjusting for positive time-of-measurement effects. The rationale was that because they are theoretically independent of particular social or environmental conditions, these estimates of adjusted change may more accurately reflect age trends in cognitive functioning. The analyses were based on published summary data from the Seattle Longitudinal Study reported in Schaie (2013), and summary data from the Betula project reported in Ronnlund and Nilsson (2008, 2009) and Ronnlund, Nyberg, Backman, and Nilsson (2005).

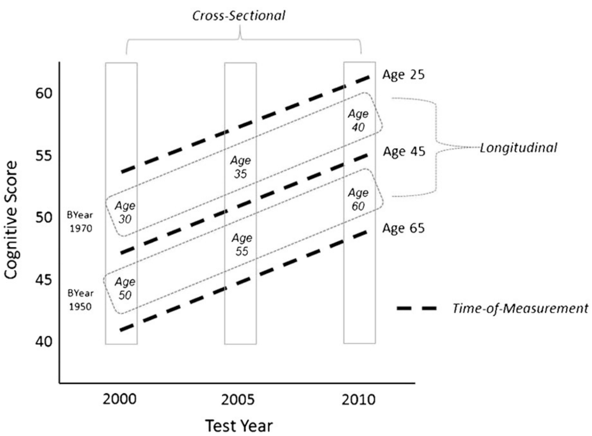

Fig. 1. Data matrix illustrating relations among time-of-measurement (Flynn Effect) influences and cross-sectional and longitudinal age comparisons.

1. Seattle Longitudinal Study (SLS)

Participants in the SLS were recruited from a Health Maintenance Organization, with similar recruitment procedures each year (Schaie, 2013, pp. 37–38). Many of the participants returned for repeat testing at 7-year intervals, and thus, the data were organized into 7-year age-groups. The cross-sectional sample with the primary cognitive battery consisted of 4,850 adults, of whom 2,777 returned for a 7-year longitudinal assessment (43% attrition). The cross-sectional sample with the latent constructs consisted of 2,038 adults (Schaie, 2013, pp. 38, 43), of whom 1,257 returned for a 7-year longitudinal assessment (38% attrition).

The same five tests were administered to new samples of adults between 22 and 77 years or older from 1956 to 1998, with between 500 and 997 new participants recruited each test year. The primary tests (i.e., series completion reasoning, spatial rotation, number arithmetic, multiple-choice vocabulary, and word fluency) were described by Schaie (2013, pp. 52–55), and all had time limits between 4 and 6 min each. Beginning in 1984, new tests were added to the assessment battery, and analyses were reported at the level of latent constructs based on factor scores across two or more tests. These tests also had time limits ranging from 1.5 to 6 min.

The cognitive scores were reported in *T*-score units (mean of 50, standard deviation of 10) based on the initial assessment of the complete sample of 4,850 across all test years for the five primary tests, and on the sample of 2,038 for the latent constructs. Data for the primary variables were obtained from Table 4.2 of Schaie (2013), and data from the latent constructs from Table 4.4. The observed 7-year longitudinal changes across all test years were obtained from Table 5.1 for the primary variables, and from Table 5.10 for the latent constructs.

2. Results

Reasoning scores of participants in four age-groups across the seven test years are portrayed in Fig. 2. Only four ages are illustrated for clarity, but the pattern was similar at all ages. Note that there were nearly parallel increases in reasoning performance as a function of test year at each age.

Regression analyses were conducted to predict the scores of each cognitive measure with age, test year, and their interaction (based on the cross-product of centered age and test year variables) as predictors. (Quadratic test-year effects were also examined, but they were not significant, and thus were not included in the final equations.)

The results of these analyses are presented in the top panel of Table 1, where it can be seen that most of the age coefficients were negative, indicating lower scores at older ages.

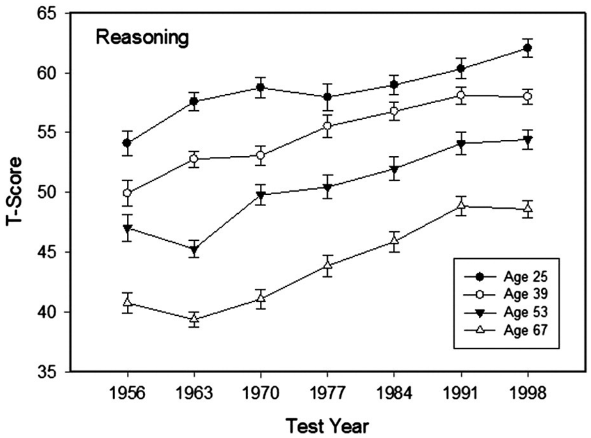

Fig. 2. Mean *T*-scores for the reasoning variable at four ages in seven different test years. Data from Table 4.2 of Schaie (2013).

The test-year coefficients were generally positive and significantly greater than zero for three of the five primary variables in the analyses with 7 test occasions across 42 years. Possibly because of the shorter interval (i.e., 3 test occasions across 14 years), the test year coefficients were not significantly different from zero for the latent constructs with the exception of verbal memory. Of particular importance in Table 1 is the absence of interactions of test year with age, indicating that there is no evidence that the time-of-measurement effects varied according to the age of the participant.

As expected from the lack of age-by-test-year interactions, Fig. 3 indicates that there were nearly parallel cross-sectional age relations for the reasoning score in the seven test years. Similar parallel trends in the other primary cognitive measures are evident in Fig. 4.5 of Schaie (2013).

The positive test-year coefficients suggest that some of the longitudinal change in cognitive performance from T1 to T2 is likely attributable to period influences on the average level of performance. Because the test-year estimate from the regression equations indicates the test score gain per year, it can be multiplied by 7 to estimate the period effect expected across the longitudinal interval used in the SLS. Adjusted longitudinal change estimates were then created by subtracting the estimated 7-year period effect from the observed longitudinal change to account for time-related gains associated with the Flynn effect. Fig. 4 portrays the 7-year period effect, and the observed and adjusted longitudinal change over this interval for the reasoning variable. Notice that the period estimates were nearly all positive and similar in magnitude across all ages, resulting in an approximately uniform decrease in the longitudinal changes at each age after the adjustment for period effects.

The period, observed longitudinal, and adjusted longitudinal values for each primary variable are presented in Table 2, and those for the latent constructs are presented in Table 3. In nearly every case, the period-adjusted longitudinal changes were more negative than the observed changes.

The values in the right-most column of the tables are averages across the different age-groups. Inspection of the entries indicates that the adjusted longitudinal changes were more negative than the observed changes when the period effects were positive. However, a few measures (e.g., number, verbal comprehension) had negative period effects, and they exhibited an opposite pattern of adjusted versus observed longitudinal change. For these measures, the observed longitudinal changes were more negative than what one would expect had there been no period effects.

Table 1
Unstandardized coefficients (with standard errors) with age, test year and their cross-product interaction term as predictors of cognitive performance (in *T*-score units).

	Age	Test Year	Interaction
Seattle Longitudinal Study Data			
Primary variables (test years from 1956 to 1998)			
Reasoning	−.38 (.01)*	.20 (.01)*	.00 (.00)
Space	−.31 (.01)*	.08 (.01)*	.00 (.00)
Verbal meaning	−.26 (.03)*	.15 (.03)*	.01 (.00)
Number	−.11 (.02)*	−.03 (.03)	.00 (.00)
Fluency	−.22 (.02)*	.01 (.02)	−.00 (.00)
Latent constructs (test years from 1984 to 1998)			
Inductive reasoning	−.40 (.02)*	.14 (.06)	−.00 (.00)
Spatial orientation	−.37 (.02)*	.09 (.06)	−.00 (.00)
Perceptual speed	−.40 (.02)*	.16 (.07)	.00 (.00)
Verbal memory	−.34 (.01)*	.11 (.04)*	−.00 (.00)
Numeric	−.06 (.03)	−.11 (.08)	.00 (.00)
Verbal comprehension	.01 (.03)	−.08 (.08)	.01 (.00)
Betula data			
Episodic memory	−.38 (.01)*	.11 (.04)*	.00 (.00)
Semantic memory	−.30 (.02)*	.15 (.04)*	.01 (.00)
Block design	−.41 (.01)*	.12 (.03)*	−.00 (.00)

* $p < .01$.

3. Betula project

Data from the Betula project were derived from cross-sectional samples tested in 1989, 1994, 1999, and 2004, and from a longitudinal sample tested in 1989 and again in 1994. The participants ranged from 35 to 80 years of age, and the recruitment procedures and inclusion criteria were described as being similar each year. There were 1000 adults in the 1989 and 1994 samples, and 500 each in the 1999 and 2004 samples. The longitudinal assessment in 1994 consisted of 875 participants (13% attrition).

An episodic memory factor was derived from scores on five tests of recall or recognition of actions or of verbal material. A semantic memory factor was derived from scores on tests of

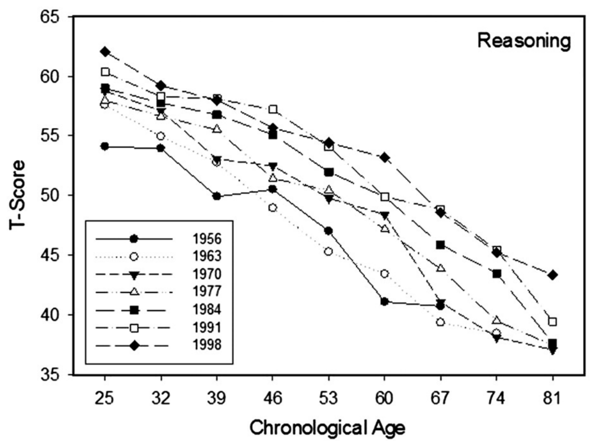

Fig. 3. Mean cross-sectional age trends for the reasoning variable (in *T*-score units) at seven test years. Data from Table 4.2 of Schaie (2013).

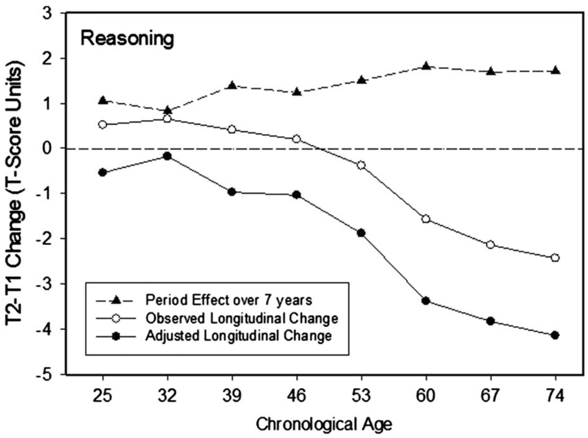

Fig. 4. Estimated time-of-measurement (period) gains over 7 years and observed and adjusted longitudinal data for the reasoning variable. Observed longitudinal data from Table 5.1 of Schaie (2013).

general knowledge, vocabulary, and three fluency tests, and the score on the Wechsler Block Design test served as a third cognitive variable.

Scores as a function of test year in Ronnlund and Nilsson (2008) were reported in z-score units, and they were converted to *T*-score units by multiplying them by 10 and adding 50. Longitudinal data were obtained from Tables 3 and 4 of Ronnlund et al. (2005), and Table 3 of Ronnlund and Nilsson (2009).

4. Results

The bottom portion of Table 1 reports results of regression analyses with age, test year, and their interaction (based on the cross-product of centered age and test year variables) as predictors of the cognitive scores. It can be seen that all age coefficients were negative, all test- year coefficients were positive, and none of the interactions was significant. Ronnlund and Nilsson (2008) reported similar results based on analyses of the raw data instead of on group means as in the current analysis.

Cross-sectional age gradients for the three cognitive scores in the four test years were portrayed in Fig. 2 of Ronnlund and Nilsson (2008). The pattern of results was similar to those in Fig. 3, and the authors noted that there was a "high similarity of age-related differences across test occasions, despite a systematic increase with regard to average performance levels (pp. 199)."

Estimates of the period effects based on comparisons between 1989 and 2004, observed longitudinal change from 1989 to 1994, and period-adjusted longitudinal changes are presented in Table 4. Note that all of the time-of-measurement estimates were positive, and that in every case the adjusted longitudinal change values were more negative than the observed changes.

5. Discussion

Consistent with the Flynn effect, the time-of-measurement effects in the analyses reported here were positive for measures of reasoning, spatial visualization, memory and speed, but were small or negative for measures of vocabulary knowledge. Importantly, the results in Fig. 2 indicates that there were nearly parallel time-of-measurement effects at four different ages, and none of the age-by-test year interactions in Table 1 was significant. There is therefore no evidence in these data that the time-of-measurement or period effects varied as a function of age.

One implication of the nearly parallel time-of-measurement effects is that cross-sectional comparisons obtained in different

Table 2
Estimates of 7-year change (in *T*-score units) in primary variables (data from 1956 to 1998) from the Seattle Longitudinal Study.

	Age								
	25	32	39	46	53	60	67	74	Average
Reasoning									
7-year period effect[a]	1.06	0.83	1.38	1.23	1.50	1.81	1.69	1.71	1.40
Obs. longitudinal[b]	0.52	0.65	0.41	0.20	−0.38	−1.57	−2.14	−2.43	−0.59
Adj. longitudinal[c]	−0.54	−0.18	−0.97	−1.03	−1.88	−3.38	−3.83	−4.14	−1.99
Space									
7-year period effect[a]	0.79	0.19	0.62	0.43	0.69	0.75	0.76	0.53	0.59
Obs. longitudinal[b]	0.99	1.44	0.43	−0.14	−0.52	−1.44	−1.90	−3.16	−0.54
Adj. longitudinal[c]	0.20	1.25	−0.19	−0.57	−1.21	−2.19	−2.66	−3.69	−1.13
Verbal meaning									
7-year period effect[a]	0.29	−0.27	0.63	0.77	1.37	1.67	1.94	2.35	1.09
Obs. longitudinal[b]	1.61	1.30	0.70	0.16	−0.37	−1.32	−2.18	−3.38	−0.44
Adj. longitudinal[c]	1.32	1.57	0.07	−0.61	−1.74	−2.99	−4.12	−5.73	−1.53
Number									
7-year period effect[a]	−0.31	−0.97	−0.70	−0.83	−0.53	−0.01	0.49	1.04	−0.23
Obs. longitudinal[b]	0.49	0.19	−0.12	−0.93	−0.72	−1.88	−2.21	−3.87	−1.13
Adj. longitudinal[c]	0.80	1.16	0.58	−0.10	−0.19	−1.87	−2.70	−4.91	−0.91
Fluency									
7-year period effect[a]	0.76	0.09	0.41	−0.25	−0.29	−0.10	0.26	0.77	0.21
Obs. longitudinal[b]	0.65	0.08	0.46	−0.39	−0.99	−1.64	−1.80	−2.68	−0.79
Adj. longitudinal[c]	−0.11	−0.01	0.05	−0.14	−0.70	−1.54	−2.06	−3.45	−1.00

[a] The test year coefficient in a regression equation predicting score from test year multiplied by 7.
[b] Obtained from Table 5.1 of Schaie (2013).
[c] Observed longitudinal change minus expected period effect.

Table 3
Estimates of 7-year change (in *T*-score units) in latent constructs (data from 1984 to 1998) from the Seattle Longitudinal Study.

	Age								
	25	32	39	46	53	60	67	74	Average
Inductive reasoning									
7-year period effect[a]	1.91	0.67	1.30	0.06	0.94	1.15	0.59	0.95	0.95
Obs. longitudinal[b]	−0.18	0.52	0.07	0.07	−0.49	−1.47	−2.03	−3.19	−0.84
Adj. longitudinal[c]	−2.09	−0.15	−1.23	0.01	−1.43	−2.62	−2.62	−4.14	−1.78
Spatial orientation									
7-year period effect[a]	1.63	0.95	0.83	−0.21	0.46	0.38	1.04	0.22	0.66
Obs. longitudinal[b]	1.30	1.01	0.09	0.35	−0.24	−1.16	−1.88	−3.72	−0.53
Adj. longitudinal[c]	−0.33	0.06	−0.74	0.56	−0.70	−1.54	−2.92	−3.94	−1.19
Perceptual speed									
7-year period effect[a]	1.65	0.85	0.34	0.31	1.09	1.27	1.20	2.06	1.10
Obs. longitudinal[b]	1.13	0.00	−0.01	−0.49	−0.69	−1.75	−2.14	−3.72	−0.96
Adj. longitudinal[c]	−0.52	−0.85	−0.35	−0.80	−1.78	−3.02	−3.34	−5.78	−2.06
Verbal memory									
7-year period effect[a]	0.32	1.26	−0.15	1.64	1.93	1.27	−0.30	0.75	0.84
Obs. longitudinal[b]	−0.63	0.55	0.02	0.02	−0.22	−1.74	−1.84	−3.53	−0.92
Adj. longitudinal[c]	−0.95	−0.71	0.17	−1.62	−2.15	−3.01	−1.54	−4.28	−1.76
Verbal comprehension									
7-year period effect[a]	−1.31	−1.21	−1.80	−1.19	−0.48	−0.50	−1.64	0.25	−0.99
Obs. longitudinal[b]	0.30	0.01	−0.58	0.43	0.07	0.27	−0.70	−1.65	−0.23
Adj. longitudinal[c]	1.61	1.22	1.22	1.62	0.55	0.77	0.94	−1.90	0.75
Numeric ability									
7-year period effect[a]	0.99	−1.67	−1.39	−2.11	−0.48	−1.17	−1.52	−1.15	−1.06
Obs. longitudinal[b]	−0.33	0.07	−0.65	−0.99	−1.34	−2.12	−2.59	−4.48	−1.55
Adj. longitudinal[c]	−1.32	1.74	0.74	1.12	−0.86	−0.95	−1.07	−3.33	−0.49

[a] The test year coefficient in a regression equation predicting score from test year multiplied by 7.
[b] Obtained from Table 5.10 of Schaie (2013).
[c] Observed longitudinal change minus expected period effect.

test years would be expected to have very similar relative age trends. This expectation was confirmed in Fig. 3, and the same pattern was evident in Fig. 4.5 of Schaie (2013), and in Fig. 2 of Ronnlund and Nilsson (2008). Nearly parallel age trends in cross-sectional comparisons of similar subtests from successive versions of the Wechsler cognitive test battery were also evident in Fig. 2.6 of Salthouse (2010). The absolute level of performance will likely be higher in more recent test years when the time-of-measurement effects are positive, but relative age comparisons appear to be similar, and equally meaningful as reflections of age–cognition relations, in each period.

In contrast to the minimal distortion of cross-sectional age–cognition trends, the existence of positive time-of-measurement effects implies that longitudinal comparisons may be distorted by the Flynn effect. Adjustments of the observed longitudinal changes for time-of-measurement effects are portrayed in Fig. 3,

Table 4
Estimates of 5-year change (in *T*-score units) from the Betula project.

	Age										
	35	40	45	50	55	60	65	70	75	80	Average
Episodic memory											
5-year period effect[a]	0.34	−0.26	0.87	0.57	0.27	0.91	0.80	0.67	0.41	0.70	0.53
Obs. Longitudinal[b]	2.00	2.08	0.77	1.43	0.78	−1.22	−0.57	−2.42	−2.63	−5.74	−0.55
Adj. Longitudinal[d]	1.67	2.34	−0.10	0.86	0.52	−2.13	−1.37	−3.09	−3.04	−6.44	−1.08
Semantic memory											
5-year period effect[a]	0.39	−0.38	1.02	0.65	0.23	1.07	1.45	0.77	0.90	1.27	0.74
Obs. Longitudinal[b]	1.70	0.65	1.76	0.61	−0.07	−0.08	−0.93	−1.61	−1.67	−2.10	−0.17
Adj. Longitudinal[d]	1.32	1.03	0.74	−0.04	−0.30	−1.15	−2.38	−2.38	−2.57	−3.37	−0.91
Block design											
5-year period effect[a]	1.27	0.94	0.87	0.10	−0.07	0.37	1.06	−0.38	0.55	1.36	0.61
Obs. Longitudinal[c]	−0.29	0.82	0.76	−0.54	−1.85	−1.51	−2.92	−2.51	−2.13	−1.77	−1.19
Adj. Longitudinal[d]	−1.56	−0.12	−0.11	−0.64	−1.79	−1.88	−3.98	−2.13	−2.68	−3.13	−1.80

Note: [a]The test-year coefficient in a regression equation predicting score from test year (1989 to 2004) multiplied by 5 to match longitudinal interval, after converting from *z*-score to *T*-score units. Time-lag estimates from Table 3 of Ronnlund and Nilsson (2008), [b]Longitudinal changes between 1980 and 1994 for episodic and semantic memory from Tables 3 and 4 of Ronnlund et al. (2005). [c]Longitudinal changes for block design from Table 3 of Ronnlund and Nilsson (2009). [d]Observed longitudinal change minus expected period effect.

and in Tables 2, 3, and 4. In nearly every case the adjusted longitudinal change was more negative than the observed change. These results imply that longitudinal decline will likely be underestimated when influences associated with positive time-of-measurement effects are ignored because each successive age is assessed at a progressively later time when test scores are generally higher.

In an earlier publication (Salthouse, 2010, Chapter 2), I have suggested that if positive period influences operate at all ages, they may operate in a manner analogous to how inflation affects salaries of nearly everyone. Specifically:

> "… some of the observed changes from one assessment period to the next may be attributable to changes in the sociocultural environment rather than to changes within the individual. That is, just as the longitudinal relation between age and salary might not be interpretable as a reflection of effects associated with aging unless adjustments are made for inflation, so might the longitudinal relation between age and cognitive performance not be meaningful until adjustments are made for historical gains in average level of performance (Salthouse, 2010, pp. 50–51)."

The results of the analyses reported here are consistent with this interpretation, and support the suggestion that longitudinal comparisons may not be meaningful without considering period or time-of-measurement influences.

The current study has a number of limitations that should be acknowledged. For example, the estimates of longitudinal change in the Betula data set were based on only the first two occasions, whereas the time-of-measurement estimates were based on four occasions. It is therefore possible that the two types of estimates are not strictly comparable because they are based on different intervals. However, it is important to note that this limitation does not apply to the comparisons based on data from the SLS, and the analyses in that data set revealed a similar pattern of results.

Another possible limitation of the analyses is that some the estimated time-of-measurement effects may be attributable to differences in selectivity or sample composition. As noted by Schaie (2013 pp. 191–192):

> "Although unlikely for large samples, it is nevertheless possible that these differences represent systematic selection effects attributable to changes in the composition of the pool from which the successive samples are drawn."

A shift in sample selectivity of this type was evident in the Virginia Cognitive Aging Project (Salthouse, 2014), as broader recruitment in more recent test years resulted in higher proportions of lower-ability individuals, and negative coefficients relating cognitive score to test year. To illustrate, for a test of vocabulary, the age-adjusted scaled scores (which have means of 10 and standard deviations of 3 in the normative sample) were 13.3 in 2001 and 10.4 in 2014. It is not clear whether sample composition shifted across test years in the SLS and Betula projects. However, the positive time-of-measurement effects in these data sets closely resemble those in studies in which the Flynn effect has been documented, and some of those were based on nationally representative samples in each test year where sample composition differences are very unlikely.

A third limitation of the current study is that although non-linear test year trends were not evident in the reported analyses, more powerful analyses may reveal non-linear time-of-measurement effects (Teasdale & Owen, 2008). Depending on the direction and magnitude of the non-linear trends, the estimates of adjusted longitudinal changes could be inaccurate for some age groups or birth cohorts.

A fourth limitation is that the conclusions are based on a number of assumptions that could be incorrect. For example, based on the lack of differential test-year effects across different ages, the Flynn effect was assumed to primarily reflect period influences rather than cohort influences. In addition, period effects were assumed to be similar in same-age different-cohort comparisons as in different-age same-cohort comparisons, which allowed longitudinal change to be adjusted for period effects. The plausibility of these assumptions, and their impact on the conclusions, should be investigated in future research.

In conclusion, the results reported here suggest that time-of-measurement effects associated with higher levels of cognitive performance at more recent times are unlikely to distort cross-sectional age trends if these effects are similar at different ages. However, because successive measurements occur at progressively more recent times in longitudinal comparisons, some of the observed change in longitudinal comparisons may be attributable to time-of-measurement effects, which when positive, will lead to underestimates of negative longitudinal change. The existence of the Flynn effect may therefore be more of a problem in longitudinal comparisons of age–cognition relations than in cross-sectional comparisons.

Acknowledgements

The project 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.

References

Baxendale, S. (2010). The Flynn effect and memory function. *Journal of Clinical and Experimental Neuropsychology*, *32*, 699–703.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, *101*, 171–191.

Hiscock, M. (2007). The Flynn effect and its relevance to neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, *29*, 514–529.

Lynn, R. (2013). Who discovered the Flynn effect? A review of early studies of the secular increase of intelligence. *Intelligence*, *41*, 765–769.

Ronnlund, M., & Nilsson, L. -G. (2008). The magnitude, generality, and determinants of Flynn effects on forms of declarative memory and visuospatial ability: Time-sequential analyses of data from a Swedish cohort study. *Intelligence*, *36*, 192–209.

Ronnlund, M., & Nilsson, L. -G. (2009). Flynn effects on sub-factors of episodic and semantic memory: Parallel gains over time and the same set of determining factors. *Neuropsychologia*, *47*, 2174–2180.

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. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Salthouse, T. A. (2010). *Major issues in cognitive aging*. New York: Oxford University Press.

Salthouse, T. A. (2014). Why are there different age relations in cross-sectional and longitudinal comparisons of cognitive functioning? *Current Directions in Psychological Science*, *23*, 252–256.

Schaie, K. W. (2013). *Developmental influences on adult intelligence* (2nd ed.). New York: Oxford University Press.

Teasdale, T. W., & Owen, D. R. (2008). Secular declines in cognitive test scores: A reversal of the Flynn effect. *Intelligence*, *36*, 121–126.

Trahan, L. H., Stuebing, K. K., Fletcher, J. M., & Hiscock, M. (2014). The Flynn effect: A meta-analysis. *Psychological Bulletin*, *140*, 1332–1360.

Williams, R. L. (2013). Overview of the Flynn effect. *Intelligence*, *41*, 753–764.