Developmental Psychology
1983, Vol. 19, No. 4, 531-543

Fourteen-Year Cohort-Sequential Analyses of Adult Intellectual Development

K. Warner Schaie and Christopher Hertzog
Division of Individual and Family Studies, Pennsylvania State University

This report provides a comprehensive analysis of data on five primary mental abilities for two longitudinal sequences (mean ages 22 to 67 years at first test) observed over two 14-year periods (Data Set 1: $N = 162$, 1956 to 1970; Data Set 2: $N = 250$, 1963 to 1977). Comparable data are also reported on cross-sectional sequences collected in 1956, 1963, 1970, and 1977. Decline in adult intelligence becomes clearly evident after age 60, with from a one-third to a one half standard deviation decrement over a 14-year period. Contrary to our earlier reports, however, small but statistically significant decrements were noted over the 53 to 67 age period, suggesting that the declines may begin prior to age 60. Cohort-sequential cohort effects, which are not attributable to confounded age changes, were found for several subtests.

Adult intellectual development has often been described in terms of a decremental model that posits continuous decline in ability after peak attainment in young adulthood. The "classic aging pattern" (Botwinick, 1977) of cross-sectional age differences in tests such as Wechsler Adult Intelligence Scale (WAIS) Block Design and Raven's Matrices but little age differences in vocabulary tests has been interpreted as an indication of decline in nonverbal intelligence over the adult life span. Horn and Cattell (1966; see also Horn & Donaldson, 1976) have interpreted this classic pattern as a reflection of different developmental trends in "fluid" and "crystallized" intelligence. In this view, fluid intelligence declines from young adulthood, which accounts for the early decrement in tests of reasoning and other skills, whereas crystallized intelligence, which represents skills and knowledge acquired through acculturation, remains stable or increases over much of the adult life span.

The validity of these qualified decremental models of adult intellectual development has been challenged by some researchers (e.g., Barton, Plemons, Willis, & Baltes, 1975; Labouvie-Vief, 1977). A major concern has been that the interpretations of the classic aging pattern summarize results from simple cross-sectional studies, which confound age changes with cohort (generational) differences (Schaie, 1965). Several years ago the first author began a study designed, in part, to disentangle the effects of age changes and cohort differences in intellectual performance by using a strategy of sequential sampling (Schaie, 1979) that consisted of repeated longitudinal and cross-sectional sampling from the same population. Earlier results (e.g., Schaie & Labouvie-Vief, 1974; Schaie, Labouvie, & Buech, 1973) were interpreted as an indication that cohort effects in intellectual performance do exist and have caused an overestimation of the magnitude and age of onset of intellectual decline (Schaie, 1979).

This interpretation of the data has been highly controversial (e.g., Baltes & Schaie, 1976; Horn & Donaldson, 1976). Indeed, previous reports by Schaie and associates have been criticized on theoretical and methodological grounds (e.g., Adam, 1978; Botwinick & Arenberg, 1976; Horn & Donaldson, 1976; Horn & McArdle, 1980). The methodological criticisms assert, at least in-

---

We are indebted to the staff and members of Group Health Cooperative of Puget Sound, who provided the subjects for the studies reported here. Data collections were supported by various grants from the National Institute on Aging.

Requests for reprints should be sent to K. Warner Schaie, Division of Individual and Family Studies, College of Human Development, Pennsylvania State University, University Park, Pennsylvania 16802.

directly, that Schaie's low estimate of age change was an artifact resulting from an inappropriate analysis of sequential designs.

This article reports a new set of analyses on Schaie's data, including results from a new data collection. Our purpose is to describe age changes in intelligence, which may be considered independent of cohort differences. We cannot fully address here the complex issues concerning the appropriate application of sequential strategies (see, however, Schaie & Hertzog, 1982). The primary problem is that the linear dependency among age, cohort, and time of measurement precludes any linear model estimating all age, cohort, and time effects and their associated interactions. Instead, only a subset of effects are estimable, conditional on assumptions that the effects not estimated are in fact nonexistent in the population. Schaie's original solution was to assume that all effects associated with one of the age, cohort, or time factors were nonexistent—resulting in a set of three sequential designs (Schaie, 1965). More recent work has shown that models making other assumptions are also estimable (e.g., Horn & McArdle, 1980). In any case, we must justify the enabling assumptions on a priori theoretical grounds, for there is no way to use the statistical results of alternative models to discover the existence of salient effects in a purely exploratory, theory-free manner.

Some of the criticisms of Schaie's early cross-sequential analyses (e.g., Schaie et al., 1973) were valid, because these studies were methodologically flawed in two important respects. First, the cross-sequential design, which crosses cohort with time of measurement, assumes that no age-associated effects are present in the data, which makes the use of cross-sequential results to evaluate age changes ill considered. Second, the early analyses compounded the problem by using unequal sampling intervals—more cohort birth years than years of measurement. This approach is not invalid per se (Schaie & Hertzog, 1982), but it does tend to produce larger proportions of variance for cohort than for time of measurement purely as a function of the difference in the number of levels for each factor. This result, however, should not have been taken as evidence that cohort accounted for more variance in intelligence than age. Moreover, confounded age effects could combine with the unequal intervals to produce cross-sequential "cohort effects," which are a function of age change alone (Botwinick & Arenberg, 1976).

If differentiation of cohort and age effects is the primary goal, then the cohort-sequential design, which crosses the age and cohort factors, is more appropriate, for it explicitly estimates age change, which is consistent across multiple cohort groups (Schaie, 1977). Since the cohort-sequential method assumes no effects of time of measurement, its cohort and age-associated parameter estimates and significance tests are internally valid only when the assumption of no time effects is true. Confounded time of measurement effects will tend to produce Age × Cohort interactions, but they will also bias the estimated age and cohort main effects. Nevertheless, we are willing to assume no time of measurement effects as a provisional hypothesis. There is no theoretical reason to expect historical period effects on intelligence test performance, although other internal validity threats (e.g., selection by period interactions) could produce time of mesurement effects in the data. Furthermore, since our main interest is in differentiating age and cohort effects, the cohort-sequential method is the preferred bifactorial design (Schaie & Baltes, 1975).[1]

Schaie and Parham (1977) reanalyzed the data from previous reports, using equal interval cohort-sequential designs on cross-sectional and longitudinal sequences.[2] Their results still suggested no major decrement in intellectual performance until after age 60.

---

[1] Horn and McArdle (1980) have criticized the use of bifactorial designs and argued for estimating a model that includes only the main effects for age, cohort, and time of measurement. The model assumes no interaction terms; in other words, it assumes additive main effects. The existence of time interaction in the population could therefore produce invalid estimates of age, cohort, and time effects. Thus, the additive model suffers from the same limitation of the more traditional bifactorial designs—if and only if the basic assumptions are met, the parameter estimates are valid (see Schaie & Hertzog, 1982).

[2] In Schaie and Parham (1977) these were termed *independent samples* and *repeated measures*, respectively. The current terminology is consistent with Schaie and Baltes's (1975) recommendations.

In 1977 another data wave was obtained from the base population, which provided us with the opportunity to extend the cohort-sequential method to estimate age changes over a 14-year interval.

Method

Subjects

The present report is part of a 21-year study of adult intellectual development; the sampling procedures have been reported in detail elsewhere (Schaie, 1979, 1983). Participants were members of a health maintenance organization (HMO) in the Seattle, Washington area. They were volunteers chosen from a randomly drawn list of the HMO membership. Testing was conducted every 7 years from 1956 to 1977. The data reported here are from two longitudinal samples and from four cross-sectional samples. The longitudinal samples included 162 participants (73 men, 89 women) originally tested in 1956 and then retested in 1963 and 1970 and 250 participants (103 men and 147 women) first tested in 1963 and then retested in 1970 and 1977. The cross-sectional samples were drawn from each testing (1956 through 1977) and covered the same birth cohorts as sampled in the longitudinal sequence. Table 1 gives the number of participants in each sex, cohort, and age group.

An important feature of the cross-sectional sampling was a switch in the population definition for the 1977 testing. On the three previous samplings, the base population was defined as the HMO membership as of 1956. By the end of the 1970 testing, this population was sufficiently "residualized" by sampling and attrition that representative sampling of older individuals in 1977 would be unlikely. Hence the population was redefined as members of the HMO at the time of testing. A collateral study suggested that the populations were quite similar but that performance levels were somewhat lower in the redefined population (Gribbin, Schaie, & Stone, Note 1).

Measurement Variables

As part of a larger battery, participants were tested with the Thurstone's 1948 Primary Mental Abilities Test (PMA, Form 11-17). In this study we report on performance on five PMA subtests: Verbal Meaning, a measure of recognition vocabulary; Space, a test requiring visualization of object rotation in two dimensions; Inductive Reasoning, a test of rule induction from alphabetic series; Number, a measure of simple addition skills; and Word Fluency, a measure of semantic retrieval based on a lexical rule.

Quasi-Experimental Design

The two longitudinal samples represent separate longitudinal sequences (Baltes, Reese, & Nesselroade, 1977), longitudinally following seven birth cohorts (partitioned in 7-year intervals) over a 14-year period. The combi-

Table 1
Cohorts, Ages, and Cell Frequencies for Cohort-Sequential Data Sets

			Sample size							
			Cross-sectional sequence						Longitudinal sequence	
Data set	Cohort[a]	Mean ages	Males[b]			Females[b]			Males	Females
1	1938	25, 32, 39	42	28	37	58	37	36	8	14
	1931	25, 32, 39	38	52	34	38	70	50	10	11
2	1931	32, 39, 46	52	34	32	70	50	37	14	26
	1924	32, 39, 46	33	71	44	37	79	43	11	15
3	1924	39, 46, 53	71	44	40	79	43	37	23	28
	1917	39, 46, 53	36	79	40	35	76	49	11	15
4	1917	46, 53, 60	79	40	35	76	49	37	25	26
	1910	46, 53, 60	35	62	38	30	81	42	17	15
5	1910	53, 60, 67	62	38	35	81	42	38	17	31
	1903	53, 60, 67	35	58	42	35	64	49	13	15
6	1902	60, 67, 74	58	42	37	64	49	33	8	10
	1896	60, 67, 74	35	64	47	37	63	41	3	12
7	1896	67, 74, 81	64	47	26	63	41	31	8	12
	1889[d]	67, 74, 81	38	38	26	38	39	24	8	6

[a] Mean birth year. [b] The *n*s here refer to independent samples of males (or females) at each data point. [c] In longitudinal sequence, mean birth year actually 1890, mean ages actually 66–80.

nation of these two longitudinal sequences yielded seven separate cohort-sequential data sets in which adjacent birth cohorts are followed over the same 14-year-age range (see Table 1). The cross-sectional sequences were similarly divided into separate cohort-sequential data sets so that cross-sectional and longitudinal analyses would be directly comparable. For each of the seven data sets, the analysis therefore involved two Age × Cohort × Sex designs, one on the cross-sectional sequences, one on the longitudinal sequences.

The data sets have unequal intervals for age and cohort. Given the 14-year-age changes, we are more likely to detect age effects, since the null hypothesis of no-age effects includes, in essence, one more 7-year level than do the cohort effects. One cannot directly compare the magnitude of F ratios or proportion of variance statistics to evaluate the relative contributions of cohort and age, but tests of the null hypothesis of no-cohort or age effects are perfectly valid in unequal interval designs.

Data Analysis

Raw scores on all subtests were scaled to T scores ($M = 50$, $SD = 10$), using the data for all cross-sectional participants through 1970 as the base population ($N = 2{,}200$).

We used multivariate analyses of variance (MANOVA) to analyze data for the five PMA subtests, using the method of hierarchical least squares (Bock, 1975). Linear and quadratic polynominal contrasts were used to represent the two degrees of freedom for age. For the analysis of cross-sectional sequences, these polynominal contrasts were treated as single degree of freedom planned comparisons. Given the nonorthogonal designs, order of effect testing is meaningful. Two models were computed: fitting age effects before cohort effects and vice versa (the sex effect was always the last main effect entered). We report cohort effects eliminating age effects and age effects eliminating cohort effects (i.e., correlated variance between effects is ignored in the analysis). The longitudinal analysis involved multivariate profile analysis for obtaining multivariate tests of the repeated measures effects of age and its interaction with cohort and sex (Bock, 1975, chapter 7).

Following a significant multivariate F, each univariate F test was examined to locate significant effects for the individual subtests. However, given the controversy regarding the existence of age decrement over the adult life span, univariate statistics are reported (and occasionally evaluated), even in the absence of multivariate significance.

Criteria for evaluation. Given a significant F test, we need to consider both the direction and magnitude of the effect. We used two criteria for this purpose: (a) the value of the effect contrast and (b) a proportion of variance statistic.[3] The sign of the contrast indicates direction. Negative linear contrasts for age indicate decline over the 14-year period, whereas negative quadratic contrasts indicate an "inverted u" or concave downward quadratic trend. In combination with a significant negative linear contrast, a negative quadratic contrast would suggest greater decline over the last 7 years of the 14-year period. The effect contrasts have been scaled to the original T-score units. A linear age contrast of −5.0, for example, represents a negative 14-year change of 5 T-score units, or half a standard deviation decline. Reference to the standard deviation gives one measure of effect size. The other measure reported here is an omega-square statistic, which estimates the proportion of variance accounted for by the effect. To adjust for the nonorthogonal design, the harmonic mean for the cell size was substituted into formulae given by Dodd and Schultz (1973) for omega-square estimates in repeated measures designs (assuming nonadditivity).

Results

Multivariate Significance Tests

Cross-sectional sequences. Significant effects for cohort, age, and sex were found in the MANOVAs for virtually all data sets (see Table 2). Cohort effects occurred for all but Data Set 4 (birth cohorts (1910 and 1917). Significant linear age effects were found throughout, except for Data Set 3 (ages 39 to 53). Data Sets 2 through 5 had significant Age × Cohort interactions. In the two youngest data sets, there was a mixture in the pattern of linear effects (some increasing, some decreasing), but beginning with the fourth data set, significant age effects reflected linear decline for at least some abilities. The cohort effects typically represented better performance for the most recently born (but this pattern was reversed for Word Fluency). Sex effects were consistent across all data sets; men performed better on Space and Number, and women excelled on Word Fluency.

Longitudinal sequences. Significant multivariate age effects were found for all data sets (see Table 3). These involved increment in performance for the first two sets (until age 46) but declined for some abilities as early as Data Set 4 (from age 46 to 60). Significant longitudinal Age × Cohort interactions occurred in Data Sets 1, 3, and 4.

Fewer cohort and sex effects were statistically reliable in the longitudinal sequences. Here, statistically significant cohort effects were found only for Word Fluency in Data Set 4, favoring the 1910 over the 1917 birth cohort. Significant sex effects corresponding to those reported for the cross-sectional se-

---

[3] We report effect contrasts instead of cell or marginal means. Detailed tables of cell means and estimates of cumulative age changes and cohort differences may be found in Schaie (1983).

Table 2
Significant Multivariate F *Tests for Cohort, Age, Sex, and the Cohort × Age Interactions: Cross-Sectional Sequences*

		Cross-sectional sequence					Cohort × Age	
Data set	Ages	*df*	Cohort	Age (linear)	Age (quadratic)	Sex	Linear	Quadratic
1	25–39	5, 504	6.30***	5.36***	2.52*	17.45***	2.23	—
2	32–46	5, 566	4.15***	3.89**	—	18.94***	3.99**	—
3	39–53	5, 613	3.89**	—	—	19.26***	2.67*	—
4	46–60	5, 588	—	2.82*	—	17.26***	4.42***	8.12***
5	53–67	5, 563	5.43***	6.26***	—	16.75***	8.70***	4.17***
6	60–74	5, 554	3.12**	13.35***	—	25.38***	—	—
7	67–81	5, 459	2.42*	8.87***	—	15.36***	—	—

* $p < .05$. ** $p < .01$. *** $p < .001$.

quences occurred only in Data Sets 1, 4, 6, and 7.

Univariate Significance Tests

Given the importance of the age-related effects for testing the decrement model, the univariate effects are considered in detail in the following secretions, which report on each of the main effects separately.

Age effects. Table 4 reports the univariate statistics for the linear and quadratic age trends. Verbal Meaning increased significantly over the period from age 25 until age 46 in the longitudinal data. In both longitudinal and cross-sectional sequences, significant decrement was first seen in the 53-to 67-year-old period; however, the significant quadratic effect indicated that the decrement is greater in the second 7 years of that period, that is, from age 60 to 67. A further acceleration of the rate of decline occurred between the ages of 74 and 81 for the longitudinal sequence. Age accounted for only a small proportion of variance for Verbal Meaning performance until age 67 (1% or less); it accounted for a maximum of 9% between ages 67 and 81 in the longitudinal data.

For Space, the longitudinal data showed a slight increase in performance from age 25 to age 39. Linear decrement initially appeared in the cross-sectional data during the 46 to 60 age range and in the longitudinal data between 53 and 67 years. The proportion of variance associated with these linear decrements again was quite small until age

Table 3
Significant Multivariate F *Tests for Cohort, Age, Sex, and the Cohort × Age Interactions: Longitudinal Sequences*

		Longitudinal sequence					
Data set	Ages	*df*[a]	Cohort	Sex	*df*[b]	Age	Cohort × Age
1	25–39	5, 35	—	3.23*	10, 30	5.15***	2.76*
2	32–46	5, 58	—	—	10, 53	2.50*	—
3	39–53	5, 69	—	—	10, 64	2.52*	3.15**
4	46–60	5, 75	2.98*	6.25***	10, 70	2.57**	2.79**
5	53–67	5, 68	—	—	10, 63	4.67***	—
6	60–74	5, 25	—	5.58***	10, 20	3.61**	—
7	67–81	5, 26	—	2.99*	10, 21	5.97***	—

[a] Degrees of freedom for the between-subjects effects of cohort and sex. [b] Degrees of freedom for the within-subjects effects of age and the Cohort × Age interaction.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 4
Linear and Quadratic Age Effects and the Associated Proportion of Variance Statistics

		Linear effects									
		Verbal Meaning		Space		Inductive Reasoning		Number		Word Fluency	
Data set (mean age)	Sequence	*d*	ω^2	*d*	ω^2	*d*	ω^2	*d*	ω^2	*d*	ω^2
1	C	1.51	—	−1.33	—	−1.39	—	2.05	.01	−1.94*	.01
(25–39)	L	2.94**	.02	1.98*	—	1.30*	.01	1.27	—	2.24*	.01
2	C	−.92	—	−1.65	—	−2.51**	.01	−1.49	—	−3.46***	.02
(32–46)	L	1.76**	.01	1.45	—	.37	—	−.24	—	−.45	—
3	C	.63	—	−1.75	—	−1.27	—	.11	—	−1.26	—
(39–53)	L	.36	—	−.17	—	.64	—	−.85*	—	−1.94	—
4	C	−1.24	—	−2.10**	.01	−1.97	—	−.53	—	−3.70**	.01
(46–60)	L	.84	—	−.04	—	−.10	—	−1.31*	.01	−2.83***	.01
5	C	−2.87*	.01	−2.39**	.01	−3.64***	.02	−2.33	—	−5.72***	.04
(53–67)	L	−1.74**	.01	−1.69*	.01	−1.55*	—	−1.90**	.01	−3.19***	.02
6	C	−5.06***	.05	−4.28***	.04	−3.32***	.03	−3.67***	.02	−7.14***	.09
(60–74)	L	−3.32**	.02	−2.76*	.01	−3.08***	.03	−2.16*	.01	−3.02***	.03
7	C	−5.83***	.07	−3.45***	.04	−2.73***	.02	−4.29***	.03	−4.30***	.03
(67–81)	L	−6.30***	.09	−5.48***	.08	−5.27***	.08	−5.33***	.07	−6.98***	.07
		Quadratic effects									
1	C	−.17	—	−1.94	—	−1.63	—	−6.03**	.02	1.61	—
(25–39)	L	−2.13	—	−1.94	—	−2.56	.01	−5.74**	.02	2.06	—
2	C	−.41	—	.65	—	.27	—	−2.73	—	3.14*	.01
(32–46)	L	−1.00	—	.04	—	−1.39	—	.42	—	2.56	—
3	C	.55	—	−1.40	—	.31	—	−1.69	—	1.87	—
(39–53)	L	−.95	—	−.43	—	1.56	—	−3.53***	.01	.10	—
4	C	1.31	—	−.57	—	2.58	—	1.38	—	3.60*	.01
(46–60)	L	−.88	—	−1.37	—	1.38	—	−.43	—	1.80	—
5	C	−4.62**	.01	−2.42*	.01	−3.27*	.01	−3.78*	.01	−2.10	—
(53–67)	L	−3.64**	.01	.05	—	−.88	—	−4.19***	.01	1.96	—
6	C	−1.76	—	.86	—	.70	—	−.57	—	3.29	—
(60–74)	L	.06	—	2.90	—	−2.07*	.01	−2.04	—	−.64	—
7	C	−.31	—	.95	—	.79	—	−.47	—	2.08	—
(67–81)	L	−4.55**	.01	1.23	—	.65	—	−3.56*	.01	1.60	—

Note. C = cross-sectional, L = longitudinal. Linear and quadratic effect contrasts, expressed in *T*-score units (.1 *SD*/unit). Negative linear contrasts indicate decline over the 14-year period; negative quadratic contrasts indicate downward concavity off the quadratic trend.
* $p < .05$. ** $p < .01$. *** $p < .001$.

67 (1% or less) and reached a maximum of 8% for the longitudinal decline over the 67 to 81 age range.

Statistically significant increment for Inductive Reasoning was found in the longitudinal data from age 25 to age 39; in contrast, in the cross-sectional sequence, significant decrement was found between ages 32 and 46. However, no significant decrement occurred for either data source in the next two data sets covering the age ranges from 39 to 53 years and from 46 to 60 years, respectively. Reliable decrement did occur for both sequences from ages 53 to 67. The significant quadratic effect for the cross-sectional data suggested that decline here was larger for the

60- to 67-year period. Another quadratic effect in the longitudinal data for Data Set 6 suggested acceleration of decrement on reasoning beyond age 67. Effects of decline as proportion of variance remained small until age 67 (2% or less) and peaked at 8% for the oldest longitudinal data set.

No significant linear increment was found for Number. However, significant quadratic effects for both data sets suggested that the young adult peak for Number was attained at about age 32. Miniscule (less than 1 T-score point) but reliable decrement was found in the longitudinal data in the 39- to 53-year-age range. The significant quadratic effect found for this age range, however, suggested that the implied decrement occurred between 46 and 53 years. A small but significant decline over the 46 to 60 age period was also found in the longitudinal sequence. By contrast, significant decrement in the cross-sectional data for Number was first found from age 60 to age 74. Significant quadratic effects for both sets during the 53- to 67-year-old period suggested greater magnitude of decrement after age 60. Despite the early detection of reliable decline for Number, the proportion of variance accounted for was 2% or less until the oldest age range where the linear decline accounted for 7% of the variance in the longitudinal data.

World Fluency was the ability with the earliest pattern of decline, at least for the cross-sectional data. For the latter, reliable decrement was observed in the youngest data set from age 25 to 39. Significant positive quadratic effects, however, suggested reversals in direction of change from age 39 to 46 and again from age 53 to 60. The longitudinal data, on the other hand, showed reliable increment to age 39, with first significant decrement in the 46- to 60-year-old period. Effect sizes remained small (less than 1%) until age 60 and peaked at 7% for the oldest longitudinal data set.

Cohort effects. Table 5 provides contrasts in T-score points, with negative values implying that the earlier born cohort performed lower than the later born cohort.

Cohort effects for Verbal Meaning favoring later born cohorts occurred only in the cross-sectional data, with reliable breaks between

Table 5

Cohort Effects in T-Score Points (1/10 Standard Deviation) and Proportion of Variance Accounted for (ω^2)

Data set (mean birth year)	Sequence	Verbal Meaning		Space		Inductive Reasoning		Number		Word Fluency	
		d	ω^2	d	ω^2	d	ω^2	d	ω^2	d	ω^2
1	C	−.34	—	−2.56***	.02	−2.65***	.03	−.18	—	.53	—
(1931, 1938)	L	−1.45	—	−4.51	.01	−4.14*	.05	−5.86*	.06	−3.14	.04
2	C	1.00	—	.26	—	−.16	—	1.90*	.01	2.95***	.02
(1924, 1931)	L	1.87	.01	.69	—	−1.15	—	−1.14	—	−1.57	—
3	C	−1.12	—	−1.17	—	−2.17**	.02	−.51	—	.94	—
(1917, 1924)	L	−1.81	—	−.52	—	−2.12	.01	−.06	—	.80	—
4	C	−.88	—	−.51	—	−.49	—	.32	—	.41	—
(1910, 1917)	L	2.38	.01	−2.12	.01	1.99	.01	1.73	—	6.18***	.08
5	C	−1.93*	.01	−2.54***	.02	−2.00**	.01	−1.75*	.01	1.07	—
(1903, 1910)	L	−1.31	—	−1.46	—	.14	—	.92	—	2.12	—
6	C	−2.00**	.01	−1.27	—	−1.77**	.01	−1.38	—	0.00	—
(1896, 1903)	L	−3.18	.02	−4.24	.02	−2.03	.02	−3.17	.01	−2.97	—
7	C	.52	—	.18	—	.41	—	−.25	—	2.45**	.02
(1889, 1896)	L	−.49	—	.10	—	2.33	—	1.87	—	.44	—

Note. C = cross-sectional, L = longitudinal.

Cohort effect contrasts expressed in *T*-score units (.1 *SD*/unit). Negative contrast indicates cohort with earlier birth year performed at a lower level than the later born cohort.

* $p < .05$. ** $p < .01$. *** $p < .001$.

the 1896, 1903, and 1910 birth cohorts. Similar findings for Space separated the 1903 from the 1910 and the 1931 from the 1938 birth cohorts. Cohort effects were most numerous for Inductive Reasoning. For both data sets, cohort effects favoring the later born cohort separated the 1931 and 1938 birth cohorts. For the cross-sectional data set, reliable cohort effects favoring the later born cohort also separated the 1896, 1903, and 1910 cohorts and the 1917 from the 1924 birth cohort. Cohort effects for Number were less consistent. In the cross-sectional data, there was a break favoring the later born between the 1903 and 1910 cohort, but a difference between the 1924 and 1931 cohort favored the earlier born. The longitudinal data had a significant effect favoring the 1938 over the 1931 cohort.

Cohort effects for Word Fluency generally ran in the opposite direction to the other abilities. All significant effects favored the earlier born cohorts. Such effects were found in the cross-sectional data between cohorts born in 1889 and 1896 and between cohorts born in 1924 and 1931, respectively. In the longitudinal data, a cohort break was observed between the 1910 and 1917 birth cohorts.

Age × Cohort interactions. In several instances the magnitude of age changes varied for the two cohorts followed over the same age range. Such interactions could imply cohort-specific differences in age changes but alternatively could also represent violations of the assumption of zero period effects (see Discussion section). Modest interactions favoring the later born cohorts occurred for Verbal Meaning in the cross-sectional data for the age ranges from 39 to 67 and for the longitudinal data for the age range from 46 to 60.[4] The only significant interaction for space favored the later born cohort in the 67- to 81-year-age range. Similar effects for Inductive Reasoning occurred for the longitudinal data for ages 53 to 74 and for the cross-sectional data from 53 to 67 years. For Number, significant interactions were found only in the cross-sectional data, again favoring later born cohorts and extending over the age range from 32 to 67 years. Most extensive interactions were found for Word Fluency. Here both data sets suggested gains for the later born cohort for the age range from 53 to 60 years but relatively greater losses for the later born cohort at ages 39 to 53. It should be noted, however, that all interaction effects accounted for 2% of the total variance or less, except for one cross-sectional interaction for Verbal Meaning at ages 39 to 53, which accounted for 4%.

Sex effects. Table 6 reports contrasts in *T* score points and proportion of variance accounted for by sex differences. Contrasts with positive signs favor the men; those with negative signs favor the women.

Although the direction of contrasts generally favored the women on Verbal Meaning and Inductive Reasoning, such differences were statistically significant only in the cross-sectional data at mean age 67 for Verbal Meaning and at mean ages 32 and 67 for Reasoning. Women on average excelled over men on Word Fluency. Such differences were statistically reliable for the cross-sectional data at all ages except 46 and 74. In the longitudinal data, however, significant differences were limited to mean ages 53, 67, and 74. By contrast, men excelled reliably on Space for the cross-sectional data at all ages and for the longitudinal data at ages 32, 46, 53, and 74. Contrasts also favored men on Number, but here effects were reliable only for the cross-sectional data at ages 39, 46, 53, and 74. Sex differences accounted for generaly less than 3% of the variance except for Space, where sex typically accounted for at least 5% of the variance (with a peak of 14% at age 32), and for Word Fluency, where sex accounted for as much as 23% at age 67.

Discussion

The results from these cohort-sequential analyses indicate that over portions of the adult life span there is age-correlated decrement in performance on the PMA independent of any generational differences. The decline becomes clearly evident after age 60 and is manifest in performance on all PMA subtests. The rate of decline after age 60 appears

---

[4] The interaction patterns were evaluated by inspection of the appropriate marginal means not reported here. Interested readers are referred to Schaie (1983) for these tables.

Table 6
Sex Effects in T-Score Points (1/10 Standard Deviation) and Proportion of Variance Accounted for (ω^2)

		Subtests									
		Verbal Meaning		Space		Inductive Reasoning		Number		Word Fluency	
Data set (mean age)	Sequence	*d*	ω^2	*d*	ω^2	*d*	ω^2	*d*	ω^2	*d*	ω^2
1	C	−1.41	.01	4.81***	.06	−1.32*	.01	1.65	—	−3.01***	.03
(32)	L	−3.46	.03	8.58**	.14	−.80	—	1.45	—	−2.10	—
2	C	−.27	—	4.50***	.07	−1.03	—	2.32**	.02	−2.55**	.01
(39)	L	−1.40	—	2.58	.02	−1.45	—	3.39	.02	−3.54	.01
3	C	.36	—	5.02***	.09	−.59	—	2.34**	.01	−.83	—
(46)	L	.99	—	3.93*	.05	.16	—	.68	—	−.89	—
4	C	−.70	—	4.56***	.07	−.36	—	1.97*	.01	−2.12**	.01
(53)	L	−2.90	.01	5.00***	.09	−2.22	.01	2.66	.01	−4.28*	.03
5	C	−1.07	—	3.88***	.06	−.75	—	1.33	—	−2.55***	.02
(60)	L	−1.33	—	.45	—	−2.24	—	−1.29	—	−3.62	.02
6	C	−3.05***	.03	3.96***	.06	−1.45*	.01	.78	—	−3.10***	.02
(67)	L	−1.90	—	1.38	—	−1.90	.01	.17	—	−8.96***	.23
7	C	−1.01	—	3.82***	.07	−.01	—	2.43**	.02	−1.83	.01
(74)	L	−.59	—	3.97*	.05	−2.12	.01	−.15	—	−6.09*	.07

Note. C = cross-sectional, L = longitudinal.
Sex effect contrasts, expressed in *T*-score units (.1 *SD*/unit). Negative contrast indicates higher performance by women.
* $p < .05$. ** $p < .01$. *** $p < .001$.

to be between one third and one half standard deviations over a 14-year interval. The age-correlated effects are quite consistent with our earlier reports that intellectual decline during the decades of the sixties and seventies is large enough to be of practical importance. However, the latest analysis of 14-year-age changes, made possible by the 1977 data collection, forces some modification of our previous position, which was that there was no reliable decline in PMA performance (with the exception of Word Fluency) prior to age 60. The present analysis revealed statistically significant decline in performance on some PMA subtests before the decade of the sixties (see Table 4). In the cross-sectional sequences, a small 14-year decline (approximately 2 T-score points) was present for Space in Data Set 4 (mean ages 46 to 60), and there was a similar, albeit nonsignificant, trend for Reasoning, In Data Set 5 (mean ages 53 to 67), all subtests showed small but mostly significant declines. However, one should not necessarily interpret the significant linear components as indicators of continuous decrements. They simply indicate statistically reliable differences between means at the two end points. In fact, both the cross-sectional and longitudinal sequences produced quadratic age components in Data Set 5, which suggests that the decline from mean age 60 to 67 was more pronounced than that from age 53 to 60. Nevertheless, the data do suggest that there are, on average, reliable but small drops in PMA performance during the fifties. Thus, our extended data indicate that the decline in PMA performance, which is pronounced after age 60, begins, on average, prior to that point.

An interesting question, and one that will probably continue to spark debate, is, "when does the average decline begin?" We expect that proponents of the early decrement model would note (a) the significant age decline for Reasoning in the cross-sectional sequence of Data Set 2 (mean ages 32 to 46) and (b) the negative direction of nonsignificant effects for Space and Reasoning in other cross-sectional sequences; therefore these proponents would argue for decline in nonverbal or fluid abilities from young adulthood (e.g., Horn & Donaldson, 1976). We have serious reservations about giving any attention to the direction of nonsignificant contrasts. But if this is done, it should be noted that the direction of the age-related contrasts for the other subtests in Date Set 2 are also negative. Is there then early decline in, say, vocabulary skills? Such a conclusion would be inconsistent with the fluid/crystallized theory as well as with the cross-sectional data it seeks to explain. A possible alternative explanation is that the population redefinition required for the 1977 testing (see the Method section) led to an overestimation of the amount of age decline in the cross-sectional sequences. Continued sampling from the redefined population will be required to address this confound.

Irrespective of the issue of age of onset, small, statistically significant declines in PMA performance during the fifties may be seen in these data. However, the results reported in Table 4 also showed reliable Cohort × Age interactions in Data Sets 4 and 5. These interactions complicate interpretation of age effects from 46 to 67 because, obviously, age-correlated changes are not uniform across the cohort pairs. When consistently observed across data sets, these interactions may indicate time of measurement effects that violate the cohort-sequential assumptions and that may affect the estimated age and cohort effects. Caution is therefore advised in interpreting the importance of these early declines, since they may be influenced in part by effects that covary with time of measurement.

The interactions for Verbal Meaning and Number consistently seem to involve greater decline in the cohort measured from 1963 to 1977, especially from 1970 to 1977 (see Footnote 4). This pattern is precisely what one would expect if the 1970 sample performed relatively high and the 1977 sample performed relatively low due to the population redefinition discussed above. The interactions for Word Fluency, on the other hand, seem to reflect greater decline between the 1956 and 1963 testings. These interactions do not appear to represent true cohort differences in aging functions. The pattern of interactions for Data Sets 2 and 3, for example, show the same pattern—the latest born cohort (measured from 1963 to 1977) shows decline over the last two intervals, whereas the other does not. Thus the pattern is inconsistent between data sets for the 1924 cohort (which is represented in both) and depends instead on the time of measurement.

Having explicitly acknowledged that reliable declines for some PMA variables occur prior to age 60, we must also emphasize that the decrement is small in magnitude, whether judged by change in standard deviation units or proportion of variance predicted. Let us not be misunderstood. We believe that the average decline prior to age 60 is statistically reliable (i.e., it does exist), but it does not appear to be of sufficient magnitude to be practically important. Changes amounting to $\frac{2}{10}$ standard deviations or less over a 14-year period, accounting for 1% or less of the variance, are not sufficiently impressive to be represented as "major"changes in intellectual performance. In this age range, individual differences in developmental trajectories are greater than the effects systematically correlated with chronological age. It is only in the 60 to 80 year age range that age declines begin to assume increasing importance.

The cohort-sequential analyses presented here do indicate that there are reliable cohort differences favoring more recently born cohorts for Reasoning, Verbal Meaning, and Space (see Table 5). These effects were generally consistent for both cross-sectional and longitudinal sequences, although they were statistically reliable only in the cross-sectional sequences, which had greater statistical power. Not all cohort pairs appeared to differ—the major break points appeared to be between the 1938 and 1931, the 1924 and 1917, and the 1910 through 1896 birth cohort pairs. These cohorts effects are unconfounded with age and clearly represent generational differences, which would contaminate age change estimates taken from simple cross-sectional data. Again, these differences are not large relative to the overall magnitude of individual differences. The cohort effects between significantly different pairs ranged between roughly 2 and 5 T-score points and usually accounted for less than 5% of the variance.

Cohort effects in the opposite direction were consistently found for Word Fluency. Again, strong Cohort × Age interactions for Word Fluency complicate these cohort differences. The pattern of the means is consistent with the view that time of measurement effects raised Word Fluency performance on the 1956 testing relative to the 1963 testing, thus producing the atypical cohort effects and the frequent Cohort × Age interactions.

As shown in Table 6, strong sex differences were found in PMA performance. Men generally performed better on Space and Number; women performed better on Word Fluency. This pattern, especially the strong sex differences on Space, is consistent with the literature on such differences in young adults (e.g., Maccoby & Jacklin, 1974). We found no indication of changes in this pattern of sex differences across the adult life span—the sex differences were quite consistent across all the data sets. The pattern of cohort and sex differences was similar between the cross-sectional and longitudinal sequences. Fewer differences were statistically reliable in the longitudinal sequences, but this appears to be simply a function of the lower statistical power to detect reliable differences between individuals.

Estimates of age effects differed between the cross-sectional and longitudinal sequences. The longitudinal sequences showed more increment in PMA performance during early adulthood and less decline until old age. The longitudinal procedure tends to produce a more select sample due to experimental mortality effects (e.g., Baltes, Schaie, & Nardi, 1971; Schaie, 1977) and affords opportunities for practice effects. However, an unpublished analysis of the larger data set suggests minimal practice effects (Gribbin & Schaie, Note 2). The more benign pattern of decline in the longitudinal sequences probably reflects less decrement in an initially more able subpopulation likely to be select with respect to health status and other variables (Riegel & Riegel, 1972). There is less intellectual decline in old subjects from this panel who have favorable life-styles (Gribbin, Schaie, & Parham, 1980) and/or no diagnosis of cardiovascular or cerebrovascular disease (Hertzog, Schaie, & Gribbin, 1978). The data from the longitudinal sequences may indicate age changes in a select subpopulation for which there appears to be relatively little decline until age 60. After that point, however, steep decline is evident in the longitudinal sequences as well.

It is interesting to note that this late life decline does not spare any of the PMA subtests. In particular, decline in Verbal Meaning performance after age 60 is of the same order of magnitude as the decline on other PMA subtests. How may this be reconciled with other data that indicate maintenance of vocabulary and other verbal skills over much of the life span? Since the PMA is a highly speeded test of intelligence, age-related decline in Verbal Meaning may represent age decline in the speed of verbal cognitive processes (e.g., speed of access and evaluation of information in semantic memory) rather than decline in verbal knowledge (vocabulary size). Indeed, it is tempting to speculate that much of the decline seen in PMA performance is a function of age changes in the speed of executing the cognitive operations required by the different subtests. For example, Berg, Hertzog, and Hunt (1982) have reported that there are age differences in mental rotation speed, as measured by the Shepard Reaction Time Task, and that mental rotation speed appears to be more highly correlated with older than with younger persons' performance on PMA Space. Cunningham and Birren (1980) have suggested that increases with age in the correlations among intelligence factors reflect the increasing importance of cognitive speed in determining adult performance on intelligence tests. Does age-related slowing in cognitive speed lead us to overestimate the loss of intellectual ability—the ability to reason, image a visual rotation, and so forth—when assessing performance on speeded tests of intelligence?

We introduced this article by discussing studies that have advanced or criticized the decremental model of adult intellectual development. The results of this study do not invalidate the decremental model, since we did find reliable age declines in cohort group means over part of the adult life span. They do indicate, however, that the cross-sectional studies cited in support of the decremental model overestimate the magnitude of mean age decline by ignoring cohort differences. Much of the disagreement about age changes in intelligence may be a function of interpretive emphasis. Others have emphasized the negative developmental slopes obtained in the most broadly representative samples as indicative of average declines in intelligence. Certainly, the data reported here do indicate decline in mean performance over part of the adult life span. On the other hand, we have emphasized the small magnitude of changes prior to old age, the large individual differences present in the data, and the limited inferences that can be drawn from average developmental curves. Age-correlated declines do not necessarily imply irreversible biological decrement, nor do small changes in averaged developmental functions imply universal decline for all individuals in the population.

We suggest that improved understanding of the effects of aging on intelligence will require less attention to average trends in global psychometric measures and more attention to (a) refinement of understanding of changes in psychometric variables, in terms of which cognitive processes appear to be contributing to the change, (b) possibility of remediation of intellectual decline, and (c) descriptive prediction of individual differences in developmental patterns.

Reference Notes

1. Gribbin, K., Schaie, K. W., & Stone, V. *Ability differences between established and redefined populations in sequential studies.* Paper presented at the meeting of the American Psychological Association, Washington, August 1977.
2. Gribbin, K., & Schaie, K. W. *Performance factors and age group differences: Practice in the face of fatigue.* Paper presented at the meeting of the Gerontological Society, Dallas, November 1978.

References

Adam, J. Sequential strategies and the separation of age, cohort, and time-of-measurement contributions to developmental data. *Psychological Bulletin,* 1978, *85,* 1309–1316.

Baltes, P., Reese, H. W., & Nesselroade, J. R. *Life-span developmental psychology: An introduction to research methods.* Monterey, Calif.: Brooks/Cole, 1977.

Baltes, P. B., & Schaie, K. W. On the plasticity of intelligence in adulthood and old age: Where Horn and Donaldson fail. *American Psychologist,* 1976, *31,* 720–725.

Baltes, P. B., Schaie, K. W., & Nardi, A. H. Age and experimental mortality in a seven-year longitudinal study of cognitive behavior. *Developmental Psychology,* 1971, *5,* 18–26.

Barton, E. M., Plemons, J. K., Willis, S. L., & Baltes, P. B. Recent findings on adult and gerontological intelligence: Changing a stereotype of decline. *American Behavioral Scientist,* 1975, *19,* 224–236.

Berg, C., Hertzog, C., & Hunt, E. Age differences in the speed of mental rotation. *Developmental Psychology,* 1982, *18,* 95–107.

Bock, D. R. *Multivariate statistical methods in behavioral research.* New York: McGraw-Hill, 1975.

Botwinick, J. Intellectual abilities. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging.* New York: Van Nostrand Reinhold, 1977.

Botwinick, J., & Arenberg, D. Disparate time spans in sequential studies of aging. *Experimental Aging Research,* 1976, *2,* 55–61.

Cunningham, W. R., & Birren, J. E. Age changes in the factor structure of intellectual abilities in adulthood and old age. *Educational and Psychological Measurement,* 1980, *40,* 271–290.

Dodd, D. H., & Schulz, R. F. Computational procedures for estimating magnitude of effects for some analysis of variance designs. *Psychological Bulletin,* 1973, *79,* 391–395.

Gribbin, K., Schaie K. W., & Parham, I. A. Complexity of life style and maintenance of intellectual abilities. *Journal of Social Issues,* 1980, *36,* 47–61.

Hertzog, C., Schaie, K. W., & Gribbin, K. Cardiovascular disease and changes in intellectual functioning from middle to old age. *Journal of Gerontology,* 1978, *33,* 872–883.

Horn, J. L., & Cattell, R. B. Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology,* 1966, *57,* 253–270.

Horn, J. L., & Donaldson, G. On the myth of intellectual decline in adulthood. *American Psychologist,* 1976, *31,* 701–719.

Horn, J. L., & McArdle, J. T. Perspectives on mathematical/statistical model building (MASMOB) in research on aging. In L. W. Poon (Ed.), *Aging in the 1980s.* Washington, D.C.: American Psychological Association, 1980.

Labouvie-Vief, G. Adult cognitive development: In search of alternative interpretations. *Merrill-Palmer Quarterly,* 1977, *23,* 227–263.

Maccoby, E. E., & Jacklin, C. N. *The psychology of sex differences.* Palo Alto, Calif.: Stanford University Press, 1974.

Riegel, K. F., & Riegel, R. M. Development, drop and death. *Developmental Psychology,* 1972, *6,* 306–319.

Schaie, K. W. A general model for the study of developmental problems. *Psychological Bulletin,* 1965, *64,* 92–107.

Schaie, K. W. Quasi-experimental designs in the psychology of aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging.* New York: Van Nostrand Reinhold, 1977.

Schaie, K. W. The Primary Mental Abilities in adulthood: An exploration in the development of psychometric intelligence. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (Vol. 2). New York: Academic Press, 1979.

Schaie, K. W. The Seattle Longitudinal Study: A twenty-one year exploration of psychometric intelligence in adulthood. In K. W. Schaie (Ed.), *Longitudinal studies of adult psychological development.* New York: Guilford, 1983.

Schaie, K. W., & Baltes, P. B. On sequential strategies in developmental research: Description or explanation. *Human Development,* 1975, *18,* 384–390.

Schaie, K. W., & Hertzog, C. Longitudinal methods. In B. B. Wolman (Ed.), *Handbook of developmental psychology.* Englewood Cliffs, N.J.: Prentice-Hall, 1982.

Schaie, K. W., & Labouvie-Vief, G. Generational versus ontogenetic components of change in adult cognitive behavior: A fourteen-year cross-sequential study. *Developmental Psychology,* 1974, *10,* 305–320.

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

Schaie, K. W., & Parham, I. A. Cohort-sequential analyses of adult intellectual development. *Developmental Psychology,* 1977, *13,* 649–653.

Received June 4, 1982
Revision received August 15, 1982 ■