International Psychogeriatrics (2017), 29:1, 137–148 © International Psychogeriatric Association 2016
doi:10.1017/S1041610216001551

Cohort effects in verbal memory function and practice effects: a population-based study

Hiroko H Dodge,[1,2] Jian Zhu,[3] Tiffany F. Hughes,[4] Beth E. Snitz,[5] Chung-Chou H. Chang,[6,7] Erin P. Jacobsen[8] and Mary Ganguli[5,8,9]

[1]*Department of Neurology, Michigan Alzheimer's Disease Center, University of Michigan, Ann Arbor, Michigan, USA*
[2]*Department of Neurology, Layton Aging and Alzheimer's Disease Center, Oregon Health & Science University, Portland, Oregon, USA*
[3]*Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, Michigan, USA*
[4]*Youngstown State University, One University Plaza, Youngstown, Ohio, USA*
[5]*Department of Neurology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA*
[6]*Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA*
[7]*Department of Biostatistics, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania, USA*
[8]*Department of Psychiatry, University of Pittsburgh School of Medicine , Pittsburgh, Pennsylvania, USA*
[9]*Department of Epidemiology, University of Pittsburgh Graduate School of Public Health , Pittsburgh, Pennsylvania, USA*

ABSTRACT

Background: In many developed countries, cognitive functioning (as measured by neuropsychological tests) appears to be improving over time in the population at large, in parallel with the declining age-specific incidence of dementia. Here, we investigated cohort effects in the age-associated trajectories of verbal memory function in older adults. We sought to determine whether they varied by decade of birth and, if so, whether the change would be explained by increasing educational attainment.

Methods: Pooling data from two prospective US population-based studies between 1987 and 2015, we identified four birth cohorts born 1902–1911, 1912–1921, 1922–1931, and 1932–1943. Among these cohorts, we compared age-associated trajectories both of performance and of practice effects on immediate and delayed recall of a 10-item Word List. We used mixed effects models, first including birth cohorts and cohort X age interaction terms, and then controlling for education and education X age interaction.

Results: We observed significant cohort effects in performance (baseline and age-associated trajectories) in both immediate recall and delayed recall, with function improving between the earliest- and latest-born cohorts. For both tests, we also observed cohort effects on practice effects with the highest levels in the latest-born cohorts. Including education in the models did not attenuate these effects.

Conclusions: In this longitudinal population study, across four decade-long birth cohorts, there were significant improvements in test performance and practice effects in verbal memory tests, not explained by education. Whether this reflects declining disease incidence or other secular trends awaits further investigation.

Key words: cognitive testing, neuropsychological testing

In many developed countries, emerging evidence suggests that cognitive functioning as measured by standardized tests is improving across generations or birth cohorts (Freedman *et al.*, 2002; Langa *et al.*, 2008; Llewellyn and Matthews, 2009; Christensen *et al.*, 2013). These findings parallel national and regional trends towards decreasing incidence rates of dementia and stroke (Qiu *et al.*, 2013; Koton *et al.*, 2014; Satizabal *et al.*, 2016). Increasing levels of education would be expected to improve cognitive reserve and slow cognitive decline (Stern, 2002); other secular trends may also play roles.

Cohort effects are variations over time, in one or more characteristics, among groups of individuals defined by some shared experience such as year or decade of birth, or a period of a specific exposure experienced at the same age during their life courses. Any given population comprises multiple sub-cohorts with different rates of exposures/outcomes, and associations between

Correspondence should be addressed to: Dr. Mary Ganguli, WPIC, 3811 O'Hara Street, Pittsburgh, PA 15213. Email gangulim@upmc.edu. Received 22 Mar 2016; revision requested 25 May 2016; revised version received 16 Aug 2016; accepted 31 Aug 2016. First published online 11 October 2016.

https://doi.org/10.1017/S1041610216001551 Published online by Cambridge University Press

them. The resulting heterogeneity within the population as a whole can mask or distort effects that might be present in its smaller, more homogeneous, constituent sub-cohorts. For example, an apparent aging effect within the population sample may in fact reflect a birth cohort effect, with the earlier-born cohorts having a higher rate of the outcome than the later-born cohorts, possibly because of specific exposures experienced during the lifetime of the earlier cohort. Yet, within each birth cohort, there may be no age effect.

From a population-based study cohort comprising four 10-year birth cohorts in southwestern Pennsylvania (USA), we have previously reported cohort effects in the cognitive domains of processing speed, language, and executive function. We observed performance improving significantly between the earliest-born and latest born cohorts, independent of the effects of increasing educational attainment (Dodge *et al.*, 2014). We now report a study to determine whether similar cohort effects would be found in verbal memory.

Methods

The data reported here are drawn from two large prospective population studies of older adults conducted in the Monongahela Valley of Southwestern Pennsylvania between 1987 and 2015. Both studies drew age-stratified random samples from the voter registration lists that are the most comprehensive publicly available listings in our region. Both studies excluded individuals who at the time of enrollment were too ill to participate, were already living in long-term care facilities, had visual or hearing impairment of sufficient severity to prevent neuropsychological testing, or had decisional incapacity. All study participants provided written informed consent at study entry and all study procedures were approved by the University of Pittsburgh Institutional Review Board for the protection of human subjects. The two studies had almost identical methods with the exceptions listed below.

1. The Monongahela Valley Independent Elders Survey (MoVIES) study recruited, from 1987 to 1989, a cohort of 1681 individuals aged 65+ from several contiguous rural communities, and followed them with biennial assessments until 2001. This study focused on the epidemiology of dementia and excluded 62 individuals who had less than sixth grade education or were not fluent in English, for whom the neuropsychological tests would not be valid. Further details have been reported previously (Ganguli *et al.*, 2000).
2. The Monongahela-Youghiogheny Healthy Aging Team (MYHAT) study recruited, from 2006 to 2008, a cohort of 1982 individuals aged 65+ from several small-town communities and follows them annually. This study focuses on the epidemiology of mild cognitive impairment and excluded 54 individuals who already had severe cognitive impairment/dementia. Further details have been reported elsewhere (Ganguli *et al.*, 2010).

Pooling data from the MoVIES and MYHAT studies, we designated each participant as having been born in one of the following four decade-long birth cohorts: 1902–1911, 1912–1921, 1922–1931, and 1932–1943. MoVIES also included 50 individuals born before 1902, whom we excluded because of their small number.

As the median educational level in both samples was high school graduate, we categorized participants into three educational attainment groups: less than high school, high school graduate, and more than high school education.

In both studies, participants underwent detailed neuropsychological assessment of cognitive domains including attention/information processing speed, executive function, language function, memory function, and visuospatial function (Ganguli *et al.*, 2010). We have previously reported cohort effects observed in the first three domains for which the two studies used identical neuropsychological tests (Dodge *et al.*, 2014). We now report cohort effects in verbal memory, doing so separately as there are two additional issues to address regarding this domain: (1) harmonization of different memory tests across the two studies, and (2) practice effects, i.e., improvements in memory test performance with repeated administration (Dodge *et al.*, 2011; Gross *et al.*, 2015).

Harmonization of memory tests across studies

In the MoVIES project, focused on identifying dementia in the population, our memory tests consisted of Immediate and Delayed Recall of a Story (Becker *et al.*, 1987) and Learning, Recall, and Delayed Recognition of the 10-item Consortium to Establish a Registry for Alzheimer Disease (CERAD) Word List (Morris *et al.*, 1989). In the later MYHAT project, focused on detecting mild cognitive impairment in the population, our memory tests included Logical Memory and Visual Reproduction from the Wechsler Memory Scale (Wechsler, 1987) and the Fuld Object Memory Evaluation Test (Fuld, 1981). Since the two studies did not use the same memory tests, it was necessary to first harmonize the memory data across the study

https://doi.org/10.1017/S1041610216001551 Published online by Cambridge University Press

populations using a standard crosswalk approach discussed later.

By the time we undertook the task, the MoVIES study had ended several years earlier and the MYHAT study was in its 9th year of annual assessments. We, therefore, administered a MoVIES verbal memory test (CERAD Word List, Immediate, and Delayed Recall) to a stratified random subsample of the MYHAT cohort, aiming to select 350 MYHAT participants (150 aged 65–74, 150 aged 75–84, and 50 aged 85+ years), approximately 50% women. Since the CERAD test was administered during the MYHAT annual assessment, the selected MYHAT participants had to perform these tests in addition to their usual MYHAT memory tests. We varied the order of test administration so that half the participants completed the Logical Memory task first and the other half completed the CERAD task first to eliminate potential order effects on test results.

Statistical analysis

Descriptive statistics: First, we examined the demographic characteristics of the pooled cohort.

Harmonizing memory test scores across studies: Next we harmonized the verbal memory test scores, employing the equipercentile equating method with log linear smoothing (Livingston, 2004) that uses percentiles to provide equivalent scores from one test to another (Monsell *et al.*, 2016). This method is simple and relies on ranking and thus does not depend on the underlying distribution of scores. Here, in a subsample of the MYHAT cohort, we performed a crosswalk analysis that converted MYHAT Logical Memory immediate and delayed recall scores into CERAD Word List immediate recall scores (sum of three trials, scored 0–30) and delayed recall scores (scored 0–10). We used the "equate" package in R (Albano, 2013) calculating 95% confidence intervals using 100 bootstrap samples (Efron, 1979).

Modeling attrition

We accounted for the possibility that the subgroup that dropped out from the study during followup might have experienced sharper age-associated cognitive declines due to comorbidities, frailty, and other reasons, resulting in biased estimates. Attrition is usually marginally associated with cognitive decline; therefore, the MCAR (missing completely at random) assumption may not hold (Little, 1995; Laird, 1988). However, if, conditional on the observed variables, attrition is not associated with cognitive decline, the missing mechanism follows MAR (missing at random), and the mixed effect models still produce valid inferences (Little, 1995). Therefore, before applying mixed effects models to compare age-associated trajectories by cohorts, we examined whether our missing data mechanism followed MAR (where missingness is associated with observed data) for both immediate and delayed recall scores. First, we modeled attrition as an outcome with subject-specific slopes (random slopes obtained from the mixed effects model) as an independent variable in the logistic regression model (model 1 in Table 3). Second, we determined whether attrition remained associated with cognitive decline after controlling for age, sex, education, and baseline cognitive scores. Model specifications are provided in the Appendix.

Modeling cohort and practice effects in memory

Having converted the MYHAT Logical Memory scores into CERAD Word List scores in an MYHAT subsample, we then fit two different linear mixed effect models (Models 1 and 2) in the entire pooled cohort. These models examined the age-associated trajectories of the CERAD immediate and delayed recall scores, and cohort effects and practice effects on the trajectories of these two outcomes. Mixed effect models adjust for within-subject correlation due to repeated measurements. The random intercepts and time (i.e., age) effects capture subject-specific deviation from the population mean intercept and age-associated trajectories. We used age as a time scale, centered at age 80 years (the average age of the all records used in the current analyses, including multiple data points per participant).

Model 1 included the following variables: fixed effects of gender, random effects of time (i.e., age centered at 80), an indicator variable for study (MoVIES vs. MYHAT), and indicator variables for each 10-year birth cohort, with the earliest-born cohort (1902–1911) as the reference group, to investigate differences across cohorts in memory scores at age 80. To assess cohort effects on age-associated trajectories of outcomes, we added Age × Birth Cohort interaction terms into the model. To examine cohort effects on practice effects, we also added interaction effects of Birth Cohort × number of Assessments (i.e., Number of times the participant took the test), using indicator variables for two to seven assessments. We included indicator variables up to the 7th assessment because preliminary analyses showed significant practice effects up to the 7th assessment. Finally, since MYHAT performed annual assessments and MoVIES assessed participants biennially, we needed to account for the possibility that

https://doi.org/10.1017/S1041610216001551 Published online by Cambridge University Press

more frequent assessments in MYHAT would amplify practice effects and differences in age-associated trajectories. Therefore, we also included an indicator variable for study (MoVIES vs. MYHAT), and interaction terms of a variable indicating MYHAT (vs. MoVIES) × age (time variable) and MYHAT × Assessment Number.

In Model 2, we further added education and Age × Education interaction terms to examine whether differences in educational attainment modified cohort effects and practice effects on the age-associated trajectories of memory scores.

Data management and analyses were performed using SAS, version 9.4 (www.sas.com) and R (http://www.gbif.org/resource/81287).

Results

The first three columns of Table 1 show the original size of each 10-year birth cohort in the two studies (MoVIES and MYHAT) and pooled across the two studies. There were 50 individuals born on or before 1901 (all in MoVIES) who were excluded for further analyses due to their small number, leaving 3613 subjects for these analyses.

Table 1 also shows the age-by-education distribution of the four birth cohorts in the pooled sample. It can be seen that the level of educational attainment increases with each successive cohort within each age category.

Table 2 shows the characteristics of the 349 randomly selected MYHAT participants who participated in the cross-walk harmonization exercise by performing both Logical Memory and CERAD Word List tasks. The mean age of this group was 75.14 (SD = 6.98). Given the smaller available pool of men in the oldest age group, the proportion of men in this age group was 45% rather than 50% as planned.

Supplementary Table 1 (available as supplementary material attached to the electronic version of this paper at www.journals.cambridge.org/jid_IPG) shows the equivalent Logical Memory and CERAD Word List scores, with 95% confidence intervals, derived from the equipercentile equating method using the MYHAT subsample ($n = 349$).

The entire pooled cohort used for the mixed effects model discussed below comprised 3613 individuals, 1631 from MoVIES and 1982 from MYHAT. The mean (SD) ages of the MoVIES and MYHAT participants were 72.43 (5.19) and 77.6 (7.44) years at baseline. Women comprised 58.1% of MoVIES and 61.04% of MYHAT. Those with less than high school graduation comprised 43.29% of MoVIES and 14.70% of MYHAT.

Table 1. Ten-year birth cohorts by studies and by age and education

STUDY	MOVIES	MYHAT	TOTAL	BASELINE AGE 65–74 YEARS			BASELINE AGE 75–84 YEARS			BASELINE AGE 85+ YEARS		
Education	n(%)	n(%)	N n(%)	<High School n(%)[a]	High School n(%)[a]	>High School n(%)[a]	<High School n(%)[a]	High School n(%)[a]	>High School n(%)[a]	<High School n(%)[a]	High School n(%)[a]	>High School n(%)[a]
Birth Decade												
1882–1901	50 (2.97)	0	50 (1.37)	*Excluded for further analyses due to small sample size*								
1902–1911	409 (24.33)	12(0.61)	421 (11.49)	0	0	0	235 (*61.4*)	58 (*15.1*)	90 (*23.5*)	24 (*63.2*)	6 (*15.8*)	8 (2.10)
1912–1921	1075 (63.95)	324 (16.35)	1399 (38.19)	358 (*37.8*)	366 (*38.7*)	223 (*23.6*)	65 (*50.4*)	33 (*25.6*)	31 (*24.0*)	91 (*28.2*)	148 (*45.8*)	84 (*26.0*)
1922–1931	147 (8.74)	928 (46.82)	1075 (29.35)	33 (*22.0*)	71 (*47.3*)	46 (*30.6*)	115 (*13.2*)	411 (47.0)	348 (*39.8*)	11 (*21.6*)	24 (*47.1*)	16 (*31.4*)
1932–1943	0	718 (36.23)	718 (19.60)	42 (*6.2*)	292 (*43.1*)	343 (*50.7*)	5 (*12.2*)	15 (*36.6*)	21 (*51.2*)	0	0	0

[a]Row percentages within each age group.

https://doi.org/10.1017/S1041610216001551 Published online by Cambridge University Press

Table 2. Demographics and memory test scores in MYHAT cross-walk sample: (N = 349)

AGE GROUPS	65–75 (N = 161)		75–85 (N = 145)		85+ (N = 43)	
	%		**%**		**%**	
Women (%)	50.30		53.10		45.95	
With less than high school education (%)	4.19		11.03		27.03	
With high school education (%)	37.13		42.76		45.95	
With more than high school education (%)	58.68		46.21		27.03	
	Mean	SD	Mean	SD	Mean	SD
CERAD word list immediate recall (sum of 3 trials)	19.66	4.60	17.38	4.96	15.65	5.10
CERAD word list delayed recall	5.17	2.55	3.77	2.55	3.32	2.53
Wechsler logical memory, immediate recall	12.17	3.89	10.61	3.53	9.62	3.59
Wechsler logical memory, delayed recall	9.43	3.93	7.72	3.81	7.41	3.70

Attrition

Over the course of followup, the cohort experienced an average mortality rate between assessment cycles ranging from 3.2% to 10.6%, and an average non-mortality dropout ranging from 4.0% to 6.5%. Details are provided in Supplementary Table 2. When we modeled attrition as an outcome with subject-specific slopes (random slopes obtained from the mixed effects model) as an independent variable in the logistic regression model (model 1 in Table 3), the result confirmed the significant association between attrition and cognitive decline ($p < 0.001$). After controlling for demographics and baseline cognitive scores, attrition was no longer significantly associated with cognitive decline (model 2 in Table 3). These results confirm that there is no evidence to reject our MAR assumption. We, therefore, applied mixed effects models to the subsequent analyses. Regarding other covariates, as expected, greater likelihood of attrition was associated with higher age at baseline, lower educational attainment, and the MoVIES study (which included more years of followup than MYHAT, which is still ongoing).

Cohort effects

Table 4 shows the results of the two linear mixed effect models (Models 1 and 2) for each of the two memory outcomes.

Immediate recall (Table 4A): Baseline scores (i.e., scores at age 80) showed steady improvement for more recently born cohorts, compared with the reference group (earliest-born cohort 1902–1911). For example, cohort 1932–1943 scored almost 6 points higher at age 80, in comparison with the reference group. CERAD immediate recall score declined by 0.43 point each year among members of the reference group. The estimated coefficients of Age × Cohort interaction terms indicated that more recently born cohort had less pronounced age-associated decline. Also, all cohorts experienced some practice effects (i.e., gains in scores) at the 2nd and subsequent assessments with the most recently born cohort consistently showing the largest gains. At and beyond the 5th assessment, the earliest-born cohort demonstrated no further practice effects, whereas the later-born cohorts continued to experience practice effects at subsequent assessment points up to 6th assessment. No gains in immediate recall scores through practice effects were observed at the 7th assessments in any cohorts. Controlling for education and Age × Cohort interactions did not change these results, although the effects of education on baseline memory scores were all significant (those with high school education and those with more than high school education had higher baseline scores than the comparison group with less than high school education). The effects of education on age-associated trajectories were also significant except for the interaction of Age × > high school education ($p = 0.13$).

Delayed recall (Table 4B): On average, delayed memory scores declined by 0.20 point per year in the reference group. Significant cohort effects were found on age-associated trajectories for Delayed Recall, as they were for Immediate Recall. The patterns of practice effects for Delayed Recall were similar to those observed for Immediate Recall

https://doi.org/10.1017/S1041610216001551 Published online by Cambridge University Press

Table 3. Assessments of missing data patterns: logistic regression with outcome being attrition, with and without including demographic variables

	MODEL 1			MODEL 2		
	ESTIMATION ODDS RATIO	95% CI	P-VALUE	ESTIMATION ODDS RATIO	95% CI	P-VALUE
	Attrition in immediate recall scores as outcome					
Subject specific slopes (random effects of age-associated slope obtained from mixed effects model)	0.03	0.02–0.05	<0.0001	0.36	0.12–1.06	0.06
Subject specific baseline (random effects of intercept obtained from mixed effects model)	–	–		0.94	0.90–0.98	0.01
Baseline age	–	–		1.10	1.07–1.14	<0.0001
Gender: female	–	–		0.64	0.55–0.75	<0.0001
Education: high school (vs. < high school)	–	–		0.91	0.75–1.11	0.64
Education: > high school (vs. < high school)	–	–		0.76	0.62–0.94	0.01
Cohort 1912–1921	–	–		0.57	0.40–0.82	0.04
Cohort 1922–1931	–	–		0.48	0.28–0.82	<0.0001
Cohort 1932–1943	–	–		0.79	0.35–1.80	0.50
Study (MoVIES)	–	–		7.59	4.52–12.77	<0.0001
	Attrition in delayed recall scores as outcomes					
Individual slopes (random effects of age-associated slope obtained from mixed effects model)	0.01	0.01–0.03	<0.0001	0.50	0.11–2.26	0.37
Individual baseline (random effects of intercept obtained from mixed effects model)	–	–		0.85	0.80–0.91	<0.0001
Baseline age	–	–		1.11	1.08–1.14	<0.0001
Gender: female	–	–		0.63	0.53–0.73	<0.0001
Education: high school (vs. < high school)	–	–		0.89	0.73–1.08	0.71
Education: > high school (vs. < high school)	–	–		0.75	0.60–0.92	0.008
Cohort 1912–1921	–	–		0.57	0.40–0.82	0.06
Cohort 1922–1931	–	–		0.47	0.27–0.79	<0.0001
Cohort 1932–1943	–	–		0.79	0.35–1.78	0.49
Study (MoVIES)	–	–		9.70	5.74–16.40	<0.0001

In Model 1, less age associated declines (i.e., larger random effects in slopes obtained from the mixed effects models) are significantly protective against attrition in both tests. However, individual specific age associated trajectory was not associated with attrition once we controlled for demographic variables. These results confirm that there is no evidence to reject our MAR assumption and mixed effects models can be applied to assess age associated cognitive trajectories. (Please see Appendix for further explanation on this issue).

except that the latest born cohort still showed practice effects at the 7th assessment. Controlling for education and for Age × Cohort interaction did not change these results. As with Immediate Recall, education was significantly associated with baseline Delayed Recall scores and age-associated trajectories over time.

Figure 1 graphically demonstrate the age-associated trajectories and practice effects by birth cohort, using the coefficients reported in the above models for Immediate Recall and Delayed Recall.

Discussion

In a population-based study of older adults comprising two pooled study samples (MoVIES and MYHAT) representing four decade-long birth cohorts, we investigated potential cohort effects

https://doi.org/10.1017/S1041610216001551 Published online by Cambridge University Press

Table 4. Cohort and practice effects on memory (linear mixed models) ($n = 3613$) 4A: immediate recall; 4B: delayed recall

	MODEL 1 COEFFICIENT	MODEL 2 P-VALUE[a]	COEFFICIENT	P-VALUE[a]
4A. Age-associated trajectories of CERAD Word List immediate recall scores				
(Intercept)	15.55	**<0.0001**	14.71	**<0.0001**
Birth cohorts (vs. Cohort 1902–1911)				
Cohort 1912–1921	0.41	0.31	0.29	0.46
Cohort 1922–1931	2.00	**0.00**	1.62	**0.00**
Cohort 1932–1943	6.10	**<0.0001**	5.52	**<0.0001**
Demographics and interactions				
Age (time variable, centered at age 80)	–0.43	**<0.0001**	–0.41	**<0.0001**
Gender: female (vs. male)	1.73	**<0.0001**	1.76	**<0.0001**
Education: high school (vs. < high school)	n/a	n/a	1.56	**<0.0001**
Education: > high school (vs. < high school)	n/a	n/a	2.51	**<0.0001**
Age × high school interaction	n/a	n/a	0.05	**0.03**
Age × >high school interaction	n/a	n/a	0.03	0.13
Study indicator: MYHAT (vs. MoVIES)[b]	–0.81	0.48	–1.34	**0.00**
MYHAT × age interaction	0.17	**<0.0001**	0.06	**<0.0001**
Age x birth cohort interactions (vs. Age × Cohort 1902–1911)				
Age × cohort 1912–1921	0.15	**0.04**	0.13	0.06
Age × cohort 1922–1931	0.29	**<0.0001**	0.27	**0.00**
Age × cohort 1932–1943	0.41	**<0.0001**	0.37	**<0.0001**
Number of assessments × cohort interactions (practice effects)				
2nd Assessment × cohort 1902–1911	0.86	**0.00**	0.77	**0.00**
2nd Assessment × cohort 1912–1921	0.39	**0.00**	0.32	**0.02**
2nd Assessment × cohort 1922–1931	0.71	**0.00**	0.66	**0.00**
2nd Assessment × cohort 1932–1943	0.81	**0.00**	0.77	**0.00**
3rd Assessment × cohort 1902–1911	1.16	**0.00**	1.00	**0.00**
3rd Assessment × cohort 1912–1921	1.19	**<0.0001**	1.07	**<0.0001**
3rd Assessment × cohort 1922–1931	1.51	**<0.0001**	1.40	**<0.0001**
3rd Assessment × cohort 1932–1943	1.80	**<0.0001**	1.70	**<0.0001**
4th Assessment × cohort 1902–1911	1.28	**0.02**	1.04	0.05
4th Assessment × cohort 1912–1921	1.63	**<0.0001**	1.45	**<0.0001**
4th Assessment × cohort 1922–1931	1.81	**<0.0001**	1.65	**<0.0001**
4th Assessment × cohort 1932–1943	2.23	**<0.0001**	2.09	**<0.0001**
5th Assessment × cohort 1902–1911	0.59	0.41	0.28	0.69
5th Assessment × cohort 1912–1921	1.91	**<0.0001**	1.67	**<0.0001**
5th Assessment × cohort 1922–1931	2.04	**<0.0001**	1.82	**<0.0001**
5th Assessment × cohort 1932–1943	2.90	**<0.0001**	2.70	**<0.0001**
6th Assessment × cohort 1902–1911	0.59	0.52	–0.04	0.96
6th Assessment × cohort 1912–1921	1.73	**<0.0001**	1.42	**0.00**
6th Assessment × cohort 1922–1931	1.73	**0.00**	1.45	**0.00**
6th Assessment × cohort 1932–1943	2.34	**<0.0001**	2.09	**<0.0001**
7th Assessment × cohort 1902–1911[b]	0	N/A	0	N/A
7th Assessment × cohort 1912–1921	–1.76	**0.00**	–1.78	**0.00**
7th Assessment × cohort 1922–1931	–0.63	**0.00**	–0.66	**0.00**
7th Assessment × Cohort 1932–1943	0.09	0.69	0.07	**0.73**

in verbal memory. Specifically, we examined performance on immediate and delayed recall of a 10-item word list and also looked at practice effects on these test scores. Our models included several interaction terms: an age × cohort interaction to identify cohort effects in the age-associated trajectories of test scores; a cohort × number of assessments interaction to address practice effects

https://doi.org/10.1017/S1041610216001551 Published online by Cambridge University Press

Table 4. Continued

	MODEL 1 COEFFICIENT	MODEL 2 P-VALUE[a]	COEFFICIENT	P-VALUE[a]
4B. Age-associated trajectories of CERAD Word List delayed recall scores				
(Intercept, treated as random effect)	4.44	**<0.0001**	4.07	**<0.0001**
Birth cohorts				
Cohort 1912–1921	–0.01	0.96	–0.06	0.77
Cohort 1922–1931	0.44	0.12	0.27	0.33
Cohort 1932–1943	2.36	**<0.0001**	2.11	**<0.0001**
Demographics and interactions				
Age (centered at age 80, treated as random effect)	–0.20	**<0.0001**	–0.19	**<0.0001**
Gender: female (vs. male)	0.74	**<0.0001**	0.75	**<0.0001**
Education: high school) (vs. < high school)	n/a	n/a	0.64	**<0.0001**
Education: > high school (vs. < high school)	n/a	n/a	1.14	**<0.0001**
Age × high school interaction	n/a	n/a	0.02	**0.04**
Age × > high school interaction	n/a	n/a	0.02	**0.04**
Study Indicator: MYHAT (vs. MoVIES)[b]	–1.79	**<0.0001**	–2.02	**<0.0001**
MYHAT × age interaction	0.10	**<0.0001**	0.09	**<0.0001**
Age × birth cohort interactions				
Age × cohort 1912–1921	0.03	0.47	0.02	0.62
Age × cohort 1922–1931	0.09	**0.02**	0.07	**0.04**
Age × cohort 1932–1943	0.16	**<0.0001**	0.15	**0.00**
Number of assessments × birth cohort interactions (practice effects)				
2nd Assessment × cohort 1902–1911	0.46	**0.00**	0.42	**0.00**
2nd Assessment × cohort 1912–1921	0.43	**<0.0001**	0.40	**<0.0001**
2nd Assessment × cohort 1922–1931	0.52	**<0.0001**	0.49	**<0.0001**
2nd Assessment × cohort 1932–1943	0.61	**<0.0001**	0.59	**<0.0001**
3rd Assessment × cohort 1902–1911	0.63	**0.00**	0.56	**0.00**
3rd Assessment × cohort 1912–1921	0.76	**<0.0001**	0.70	**<0.0001**
3rd Assessment × cohort 1922–1931	0.92	**<0.0001**	0.88	**<0.0001**
3rd Assessment × cohort 1932–1943	1.19	**<0.0001**	1.15	**<0.0001**
4th Assessment × cohort 1902–1911	0.57	**0.04**	0.46	0.09
4th Assessment × cohort 1912–1921	0.93	**<0.0001**	0.85	**<0.0001**
4th Assessment × cohort 1922–1931	1.12	**<0.0001**	1.05	**<0.0001**
4th Assessment × cohort 1932–1943	1.34	**<0.0001**	1.28	**<0.0001**
5th Assessment × cohort 1902–1911	0.34	0.36	0.20	0.59
5th Assessment × cohort 1912–1921	1.11	**<0.0001**	1.01	**<0.0001**
5th Assessment × cohort 1922–1931	1.23	**<0.0001**	1.14	**<0.0001**
5th Assessment × cohort 1932–1943	1.65	**<0.0001**	1.57	**<0.0001**
6th Assessment × cohort 1902–1911	–0.02	0.97	–0.19	0.69
6th Assessment × cohort 1912–1921	1.04	**<0.0001**	0.91	**<0.0001**
6th Assessment × cohort 1922–1931	1.17	**<0.0001**	1.06	**<0.0001**
6th Assessment × cohort 1932–1943	1.69	**<0.0001**	1.58	**<0.0001**
7th Assessment × cohort 1902–1911	0	N/A[b]	0	N/A[b]
7th Assessment × cohort 1912–1921	–0.01	0.96	–0.02	0.92
7th Assessment × cohort 1922–1931	–0.01	0.94	–0.02	0.83
7th Assessment × cohort 1932–1943	0.45	**<0.0001**	0.44	**<0.0001**

MYHAT = Monongahela-Youghiogheny Healthy Aging Team; MoVIES = Monongahela Valley Independent Elders Survey
Interactions of MYHAT and 2nd, 3rd, …, up to 7th assessment were also controlled in the models (not shown) to factor in the different assessment schedules (annual vs. biennial) in the two studies.
[a]Bold values indicate statistical significance.
[b]No subject had seven or more assessments in this birth cohort.

https://doi.org/10.1017/S1041610216001551 Published online by Cambridge University Press

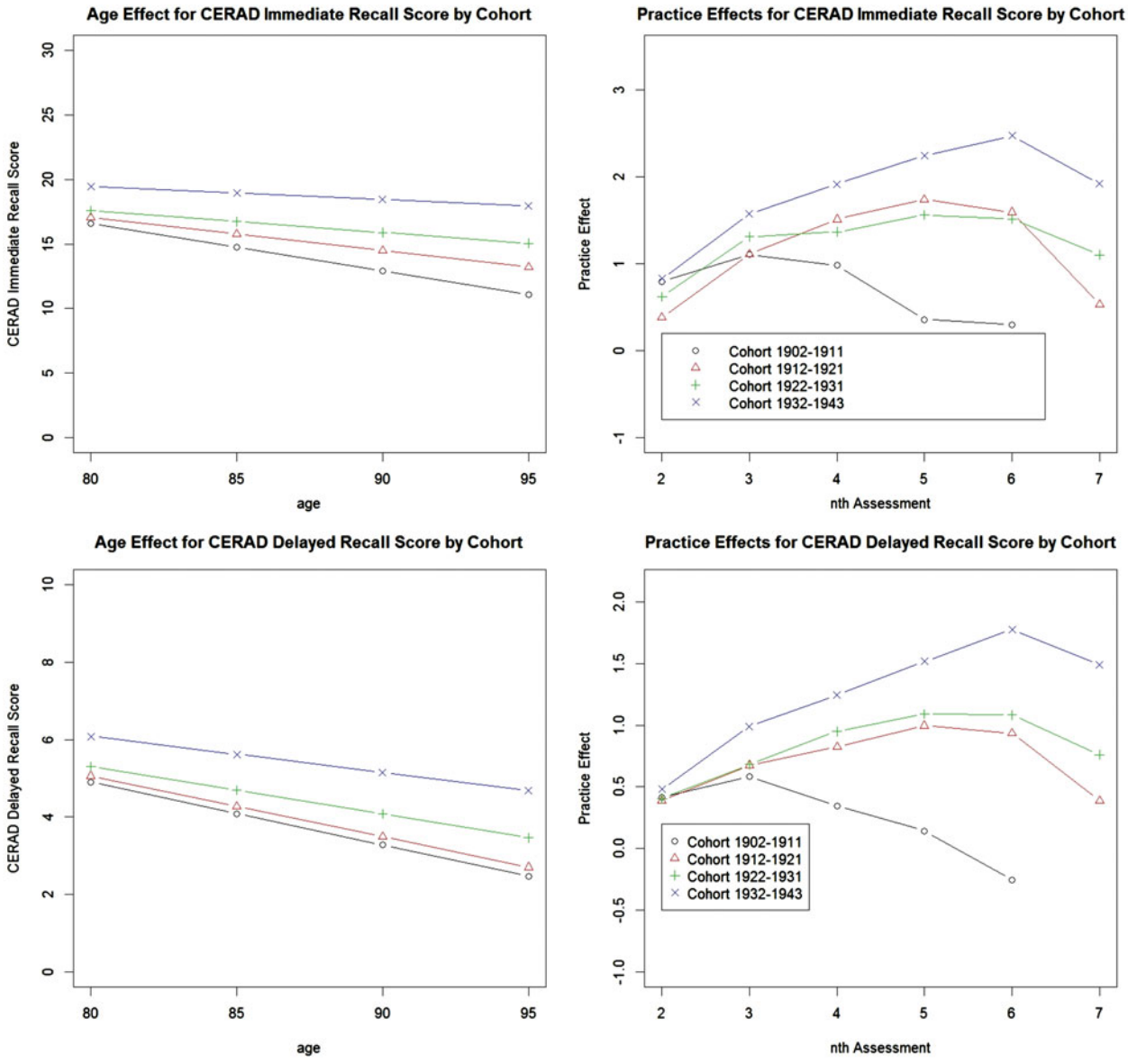


Figure 1. Cohort and practice effects on age-associated memory trajectories of immediate and delayed recall.

with greater numbers of repeated assessments; an age × education interaction term to factor in the effect of education; a study × age interaction term and a study × assessment number interaction term to account for design differences between the two studies, specifically, the fact that assessments were annual in MYHAT and biennial in MoVIES. We found significant cohort effects in baseline scores, age-associated trajectories, as well as practice effects favoring the latest-born cohorts, on both immediate and delayed recall scores. Although education strongly influenced baseline and the trajectories of test scores over time on both tests, none of the observed cohort effects were explained by level of education.

The importance of detecting birth cohort effects (characteristics restricted to a group of people born at the same time) is in part to distinguish them from age effects (characteristics associated with aging regardless of when people are born) and from period effects (characteristics associated with living through a specific historical period of time, perhaps linked to an exposure that occurred only during that time). Identification of a true cohort effect in cognition can shed light on potential underlying mechanisms related to secular trends, whether in changes in disease incidence or in differential exposures over that period. One example, which we did investigate here, is years of education. Other examples might include changes in risk factors,

https://doi.org/10.1017/S1041610216001551 Published online by Cambridge University Press

such as smoking and obesity, or the introduction of effective therapies for cardiovascular disease.

Previous studies examining cohort effects in cognition have included the US nationwide studies HRS and AHEAD, which found significant decreases in the proportion of individuals with severe cognitive impairment between 1993 and 1998 (Freedman *et al.*, 2002) and between 1993 and 2002 (Langa *et al.*, 2008). A British national study found an increase in semantic verbal fluency between 1991 and 2002 (Llewellyn and Matthews, 2009). A Danish national study compared birth cohorts born in 1905 and 1915, and found the later born cohort performed significantly better on the Mini-Mental State Examination (MMSE) and on a composite of five cognitive tests (Christensen *et al.*, 2013). A Swedish study showed better performance at age 70 in a cohort born in 1930 compared to a cohort born in 1901–1902 (Sacuiu *et al.*, 2010). We ourselves previously reported from the same pooled cohort that there were significant cohort effects favoring the more recently born cohorts on tests tapping the domains of processing speed, executive function, and language (Dodge *et al.*, 2014).

A well-established phenomenon, called the Flynn Effect (Flynn, 1987) refers to rising IQ levels in successive generations of young people, thus requiring IQ tests to be re-standardized periodically. This effect has been variously attributed to secular trends in factors such as improved nutrition, healthcare, education, and environmental stimulation, possibly contributing to "brain reserve" as well as "cognitive reserve" (Osone *et al.*, 2016). Such early-life influences might play a potential role in affecting age-associated trajectories in older adults.

Our current report is focused on birth cohort effects in age-associated cognitive trajectories rather than on prevalence or incidence of age-related diseases. However, the underlying presence of disease is a potential explanation for the improving trends we have observed in memory test performance as reported here and in other cognitive domains as we reported earlier. There is a growing literature on trends in the prevalence and incidence of dementia (Schrijvers *et al.*, 2012; Matthews *et al.*, 2013; Satizabal *et al.*, 2016), a leading cause of cognitive impairment in older adults. Since prevalence is a function of both incidence and duration of disease, these data need to be further evaluated in the context of mortality trends; most studies show increasing survival favoring recent cohorts.

Secular trends in risk factors for dementia should be investigated as potential reasons for the favorable trends in dementia incidence and in cognitive functioning. As we lack lifetime data on our participants, we can only speculate that the improving cohort effects might be related to factors suggested in the literature, such as declines in smoking (CDC), stroke (Koton *et al.*, 2014), and poverty (O'Brien *et al.*, 2010), and increases in education and exposure to information technology (Bordone *et al.*, 2015). On the other hand, there are increasing trends in obesity (Flegal *et al.*, 2012) and diabetes (Gregg *et al.*, 2014), which might not bode as well for the next generation of older adults.

Our study was made possible by our having access to data from two studies that were carried out similarly in contiguous areas over different periods. For reasons related to their primary objectives, the two studies used different memory tests that had to be harmonized for the present analyses. Since the first study was over by the time the second one began, a random subsample of the second study had to take an additional test that was a part of the first study. Had we the resources to test a larger subsample, the results of the cross-walk analysis might have been more robust and allowed us to equate scores within age or educational subgroups rather in the subsample as a whole. Whether other minor differences between the designs of the two studies, listed under Methods, may potentially have affected the overall results is not knowable.

Future research should include investigating potential explanatory factors for these population trends, ideally in life-course studies that examine exposures across the lifespan. These findings should also be replicated among populations of low and middle income countries, and in racial/ethnic minorities in the more affluent countries. These trends can potentially shed light on underlying mechanisms as well as facilitate future planning and policymaking as the world's population ages.

Conflicts of interest

None

Description of authors' roles

Dr. Dodge designed the study, performed the analyses, and co-wrote the manuscript. Mr. Zhu assisted Dr. Dodge with the analyses. Drs. Snitz, Hughes, Chang, and Ms. Jacobsen participated in interpretation and reporting of the results and literature review and critical revisions to the manuscript. Dr. Ganguli co-designed the study and took primary responsibility for writing the manuscript.

https://doi.org/10.1017/S1041610216001551 Published online by Cambridge University Press

Acknowledgments

The authors acknowledge the efforts of both MoVIES and MYHAT study teams, and the contributions of the senior citizens who participated in both studies. The work reported here was supported in part by grants # R01AG07562, R01AG023651, K07AG044395, P30AG008017 from the National Institute on Aging, National Institutes of Health, United States Department of Health and Human Services.

Supplementary material

Supplementary Table S1 is available as supplementary material attached to the electronic version of this paper at http://dx.doi.org/10.1017/S1041610216001551

References

Albano, A. (2013). *Equate: Statistical methods for test score equating. R package Version 1.2-0*. Available at: http://CRAN.R-project.org/package=equate.

Becker, J. T., Boller, F., Saxton, J. and McGonigle-Gibson, K. L. (1987). Normal rates of forgetting of verbal and non-verbal material in Alzheimer's disease. *Cortex*, 23, 59–72.

Bordone, V., Scherbov, S. and Steiber, N. (2015). Smarter every day: the deceleration of population ageing in terms of cognition. *Intelligence*, 52, 90–96.

Centers for Disease Control and Prevention. (2016). *Trends in Current Cigarette Smoking Among High School Students and Adults, United States, 1965-2014. National Health Interview Survey, 1965-2014*. Available at: http://www.cdc.gov/tobacco/data_statistics/tables/trends/cig_smoking/index.htm; last accessed 2016.

Christensen, K. *et al.* (2013). Physical and cognitive functioning of people older than 90 years: a comparison of two Danish cohorts born 10 years apart. *Lancet*, 382, 1507–13.

Dodge, H. H., Wang, C. N., Chang, C. C. and Ganguli, M. (2011). Terminal decline and practice effects in older adults without dementia: the MoVIES project. *Neurology*, 77, 722–30.

Dodge, H. H., Zhu, J., Lee, C. W., Chang, C. C. and Ganguli, M. (2014). Cohort effects in age-associated cognitive trajectories. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 69, 687–94.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics*, 7, 1–26.

Flegal, K. M., Carroll, M. D., Kit, B. K. and Ogden, C. L. (2012). Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. *JAMA*, 307, 491–7.

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

Freedman, V. A., Aykan, H. and Martin, L. G. (2002). Another look at aggregate changes in severe cognitive impairment: further investigation into the cumulative effects of three survey design issues. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 57, S126–31.

Fuld, P. A. (1981). *Fuld Object-Memory Evaluation*. Woodale, IL: Stoelting Co.

Ganguli, M., Dodge, H. H., Chen, P., Belle, S. and DeKosky, S. T. (2000). Ten-year incidence of dementia in a rural elderly US community population: the MoVIES Project. *Neurology*, 54, 1109–16.

Ganguli, M., Snitz, B. E., Lee, C. W., Vanderbilt, J., Saxton, J. A. and Chang, C. C. (2010). Age and education effects and norms on a cognitive test battery from a population-based cohort: the monongahela-youghiogheny healthy aging team. *Aging and Mental Health*, 14, 100–7.

Gregg, E. W. *et al.* (2014). Changes in diabetes-related complications in the United States, 1990-2010. *New England Journal of Medicine*, 370, 1514–23.

Gross, A. L. *et al.* (2015). Predictors of retest effects in a longitudinal study of cognitive aging in a diverse community-based sample. *Journal of the International Neuropsychological Society*, 21, 506–18.

Koton, S. *et al.* (2014). Stroke incidence and mortality trends in US communities, 1987 to 2011. *JAMA*, 312, 259–68.

Laird, N. M. (1988). Missing data in longitudinal studies. *Statistics in Medicine*, 7, 305–15.

Langa, K. M. *et al.* (2008). Trends in the prevalence and mortality of cognitive impairment in the United States: is there evidence of a compression of cognitive morbidity?. *Alzheimer's & Dementia*, 4, 134–44.

Little, R. J. A. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association*, 90, 1112–21

Livingston, S. A. (2004). *Equating Test Scores (without IRT)*. Princeton, NJ: ETS.

Llewellyn, D. J. and Matthews, F. E. (2009). Increasing levels of semantic verbal fluency in elderly English adults. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, 16, 433–45.

Matthews, F. E. *et al.* (2013). A two-decade comparison of prevalence of dementia in individuals aged 65 years and older from three geographical areas of England: results of the cognitive function and ageing study I and II. *Lancet*, 382, 1405–12.

Monsell, S. E. *et al.* (2016). Results from the NACC uniform data set neuropsychological battery crosswalk study. *Alzheimer Disease and Associated Disorders*, 30, 134–39.

Morris, J. C. *et al.* (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, 39, 1159–65.

O'Brien, E., Wu, K. B. and Baer, D. (2010). *Older Americans in Poverty: A Snapshot*. Washington, DC: AARP Public Policy Institute.

Osone, A., Arai, R., Hakamada, R. and Shimoda, K. (2016). Cognitive and brain reserve in conversion and reversion in patients with mild cognitive impairment over

https://doi.org/10.1017/S1041610216001551 Published online by Cambridge University Press

12 months of follow-up. *Journal of Clinical and Experimental Neuropsychology* Epublished ahead of print. doi: 10.1080/13803395.2016.1191620.

Qiu, C., von Strauss, E., Backman, L., Winblad, B. and Fratiglioni, L. (2013). Twenty-year changes in dementia occurrence suggest decreasing incidence in central Stockholm, Sweden. *Neurology*, 80, 1888–94.

Sacuiu, S. *et al.* (2010). Secular changes in cognitive predictors of dementia and mortality in 70-year-olds. *Neurology*, 75, 779–85.

Satizabal, C. L., Beiser, A. S., Chouraki, V., Chene, G., Dufouil, C. and Seshadri, S. (2016). Incidence of dementia over three decades in the Framingham heart study. *New England Journal of Medicine*, 374, 523–32.

Schrijvers, E. M., Verhaaren, B. F., Koudstaal, P. J., Hofman, A., Ikram, M. A. and Breteler, M. M. (2012). Is dementia incidence declining?: Trends in dementia incidence since 1990 in the Rotterdam study. *Neurology*, 78, 1456–63.

Stern, Y. (2002). What is cognitive reserve? Theory and research applications of the reserve concept. *Journal of the International Neuropsychological Association*, 8, 448–460.

Wechsler, D. (1987). *Wechsler Memory Scale Revised*. San Antonio, TX: The Psychological Corporation.

APPENDIX

Additional model descriptions used for Table 3.

Linear mixed model to obtain unadjusted baseline cognitive score and cognitive decline for both cognitive scores $Y_{it} = \beta_0 + \beta_{i0} + (\beta_1 + \beta_{i1})t + \beta_0 + \varepsilon_{it}$, where $(\beta_{i0}, \beta_{i1}) \sim BVN(\mu, \Sigma)$ and $\varepsilon_{it} \sim N(0, \sigma^2)$.

The estimated $\hat{\beta}_{i0}$ and $\hat{\beta}_{i1}$ are then used as subject specific baseline and cognitive decline.

For both cognitive scores, model 1 examines the marginal association between attrition and cognitive decline: logit(Pr(subject i is missing due to attrition)) $= a_0 + a_1\hat{\beta}_{i1}$.

Model 2 examines the association between attrition and cognitive decline after controlling for the fully observed variables logit(Pr(subject i is missing due to attrition)) $= a_0 + a_1 + \hat{\beta}_{i1} + a_2\ \hat{\beta}_{i0} + a_3$ baseline age$+a_4$ gender$+a_5$ education$+\boldsymbol{a}_6$ cohort$+a_7$ study indicator.

https://doi.org/10.1017/S1041610216001551 Published online by Cambridge University Press