

Research Article

Using Mixture Modeling to Construct Subgroups of Cognitive Aging in the Wisconsin Longitudinal Study

Sara M. Moorman, PhD,^{1,*} Emily A. Greenfield, PhD,^{2,} and Kyle Carr, MA^{1,}

¹Boston College, Chestnut Hill, Massachusetts. ²Rutgers, The State University of New Jersey, New Brunswick.

*Address correspondence to: Sara M. Moorman, PhD, Department of Sociology, Boston College, 140 Commonwealth Avenue, McGuinn Hall 426, Chestnut Hill, MA 02467. E-mail: moormans@bc.edu

Received: June 4, 2020; Editorial Decision Date: October 20, 2020

Decision Editor: Angela Gutches, PhD

Abstract

Objectives: Longitudinal surveys of older adults increasingly incorporate assessments of cognitive performance. However, very few studies have used mixture modeling techniques to describe cognitive aging, identifying subgroups of people who display similar patterns of performance across discrete cognitive functions. We employ this approach to advance empirical evidence concerning interindividual variability and intraindividual change in patterns of cognitive aging.

Method: We drew upon data from 3,713 participants in the Wisconsin Longitudinal Study (WLS). We used latent class analysis to generate subgroups of cognitive aging based on assessments of verbal fluency and episodic memory at ages 65 and 72. We also employed latent transition analysis to identify how individual participants moved between subgroups over the 7-year period.

Results: There were 4 subgroups at each point in time. Approximately 3 quarters of the sample demonstrated continuity in the qualitative type of profile between ages 65 and 72, with 17.9% of the sample in a profile with sustained overall low performance at both ages 65 and 72. An additional 18.7% of participants made subgroup transitions indicating marked decline in episodic memory.

Discussion: Results demonstrate the utility of using mixture modeling to identify qualitatively and quantitatively distinct subgroups of cognitive aging among older adults. We discuss the implications of these results for the continued use of population health data to advance research on cognitive aging.

Keywords: Alzheimer's disease and related dementias, Episodic memory, Life course, Mixture models, Verbal fluency

On average, individuals experience declines in cognition as they advance through adulthood (Salthouse, 2010). Studies based on large population samples have found small, yet robust, age-associated losses in particular domains of cognition even as early as one's 30s (Hughes et al., 2018). Thus, one approach has been to focus on mean differences across age groups, to uncover theoretically universal processes of senescence (Hertzog, 2008). From this perspective, within-cohort variation in cognition is due primarily to differences in overall intellectual ability (Lindenberger, 2014).

In contrast, an individual differences approach to the field holds that differences in cognitive performance—both

in terms of differences between persons in performance at a single point in time, as well as differences in continuity and change within individuals' performance as they age—reflect “meaningful variation” involving complex biological, psychological, and social processes (Salthouse, 2017, p. 7). As Hertzog (2008) stated: “Although there are normative changes across the adult life span at biological, psychological, and social levels, there is also diversity in the expression of age-related changes in structures and mechanisms on cognition” (p. 34). In short, cognitive aging might not be a singular entity, but instead, be characterized by qualitatively distinct trajectories of performance

across multiple domains of cognition (Casaletto et al., 2019).

Our study contributes to calls to account for individual differences in patterns of cognitive aging (Hertzog, 2008) by examining the utility of mixture modeling to distinguish subgroups of cognitive aging. We apply mixture modeling to data from cognitive assessments with participants in the Wisconsin Longitudinal Study (WLS), which is among the longest-running prospective cohort studies in the United States. We first use latent class analysis (LCA) to identify discrete subgroups of cognitive aging when participants are age 65 and 72, and then we employ latent transition analysis (LTA) to examine interindividual differences in intraindividual change between ages 65 and 72.

Conceptual Approaches to Heterogeneity in Later-Life Cognition

The extent of individual differences in cognition increases across the life course, such that the cognition of two randomly selected 5-year-olds will be more similar than the cognition of two randomly selected 75-year-olds (Lindenberger, 2014). Thus, experimental psychologists have advanced theoretical concepts to delineate categories of variability in human cognition (MacDonald et al., 2009). “Diversity” refers to varied performance levels between persons (Gorus et al., 2006). There are multiple terms for discussing within-person differences, and this study concerns two: “dispersion” and “intraindividual change.” Dispersion refers to an individual’s differential performance levels across neurocognitive domains or tasks (Costa et al., 2019). Intraindividual change, then, refers to variability within persons across occasions that is both long term and enduring (Nesselroade, 1991). In the present study, we document individual differences, or diversity, in dispersion and in intraindividual change.

Dispersion appears to be an indicator of cognitive robustness or reliability, where fluctuating performance across domains can be maladaptive (Costa et al., 2019). Most studies of dispersion calculate an index value for each participant, and do not investigate the specific domains in which participants score well or poorly. The curve of the dispersion index across the life course is U-shaped, with both children and older adults showing the highest levels (Mella et al., 2016). Dispersion is also higher on tasks of greater number and complexity (Gorus et al., 2006), as well as among people in the early stages of Alzheimer’s disease and related dementias (ADRD) (Costa et al., 2019; Roalf et al., 2016).

Intraindividual change in cognitive function can be an indicator of either normative development or disease-related processes. Being unmarried, being a member of a marginalized racial/ethnic group, and being of low socioeconomic status are factors that are associated with lower peak performance in adulthood as well as more precipitous decline with age (Karlman et al., 2009). These factors

likely mark more proximal causes, such as neighborhood disorder, environmental toxins, and psychological distress (Sharifian et al., 2020). The most common statistical tool to study intraindividual change in cognitive aging is latent growth curve models; however, recent methodological research suggests that these models are prone to being misspecified and misinterpreted within studies of cognitive aging (Ghisletta et al., 2020). These developments accentuate the need for a range of methodological approaches to studying both dispersion and intraindividual change in later-life cognition.

Mixture Models and Subgroups of Cognitive Aging

Scholars approaching the study of cognitive aging from a life span developmental approach have called for the greater use of person-centered methodologies to model interindividual differences in intraindividual variability, thereby “linking within-person changes to more traditional dimensions of individual differences” (Hertzog, 2008, p. 38). Mixture modeling is one such idiographic approach, as it is a statistical method designed to compare and contrast groups of individuals across a set of variables rather than to assess relationships among variables (B. Muthén & L. K. Muthén, 2000). Mixture modeling takes a probabilistic approach to describe population subgroups that are qualitatively different from each other across a variety of types and degrees of characteristics (Collins & Lanza, 2010). Although mixture modeling has been widely used in the field of adolescent development, its use in the field of cognitive aging has been more limited. We review the limited number of studies in this area below.

Costa and colleagues (2013) first used mixture modeling with data from 506 community-dwelling Portuguese older adults, although their results showed quantitative groups of low, medium, and high scorers across domains, rather than qualitative groups with domain-specific performance levels. Their results thus echoed those of other studies that have used more traditional statistical methods to demonstrate significant diversity in quantitative levels of cognitive performance (e.g., Goh et al., 2012).

However, other studies employing mixture models have yielded evidence for qualitatively distinct subgroups of cognitive aging. Zammit and colleagues have used the approach most extensively with data from the Rush Memory and Aging Project (Zammit et al., 2020; Zammit, Hall, et al., 2019; Zammit, Muniz-Terrera, et al., 2019). The Rush study measured multiple domains of cognition among 1,662 Chicago-area older adults. LCAs revealed five cognitive subgroups cross-sectionally: groups that scored poorly, average, and highly across tests, as well as a subgroup of people showing memory impairment and a subgroup of people showing perceptual impairment. Over 3 years, surviving participants who transitioned away from their baseline subgroup were at the greatest risk of receiving a

diagnosis of dementia. Among participants who died, the members of the low scoring and memory impairment subgroups showed the most Alzheimer's disease pathology on autopsy.

Zammit and colleagues also have applied mixture models to data from the Einstein Aging Study (Zammit, Hall, Katz, et al., 2018; Zammit, Hall, Lipton, et al., 2018). They identified cognitive subgroups among 1,345 older adults in New York City. Individuals with overall low scores and those with memory impairments were most likely to receive a dementia diagnosis within 4 years. Finally, other researchers' analysis of 6 years of data from participants aged 78 and older in the Health and Retirement Study (HRS) identified three subgroups of cognition: normal cognitive function, fluid intelligence impairment, and cognitive impairment (Huang et al., 2019). They found that once classified as cognitively impaired, no participants transitioned over the 6 years to normal function or fluid intelligence impairment only.

Focus of the Current Study

In the present study, we use mixture models to advance empirical evidence for subgroups of cognitive aging, as well as how individuals transition across subgroups over a 7-year period. We use data from the WLS, which is one of the longest-running and largest cohort studies in the United States. The WLS allows for a multidimensional analysis of later-life cognition at ages 65 and 72 given its inclusion of assessments of multiple cognitive functions. Specifically, we aim to describe subgroups of cognitive aging at ages 65 and 72—orienting to whether the subgroups within the population at large remain stable or change. We also aim to describe the extent to which WLS participants make transitions between particular subgroups between the ages of 65 and 72.

Method

Data

The WLS is a random sample of one-third of the men and women who graduated from Wisconsin high schools, both public and private, in 1957 ($N = 10,317$). Respondents were first surveyed during their senior year of high school in 1957 and then contacted again in 1975 (36 years old), 1993 (54 years old), 2004 (65 years old), and 2011 (72 years old). For the purposes of this study, we focus on the waves from 2004 and 2011 when individuals were 65 and 72 years old, respectively, and when comprehensive cognitive batteries appeared in the WLS.

In 1960, 97% of Wisconsin residents were white (Wisconsin Legislative Reference Bureau, 2017). Thus in the WLS, no variable indicates participants of color for reasons of privacy and confidentiality, and there are too few in number for statistical analysis (Herd et al., 2014).

Thus, because of racial/ethnic and educational homogeneity, as well as the study's regional focus, the results of analyses of WLS data cannot and should not be generalized to the general population of U.S. older adults.

Our analytic sample excluded several groups from the original sample of 10,317 participants. [Supplementary Table 1A](#) shows the process of selection of the analytic sample. First, 3,052 participants left the study before age 65 because of death ($n = 1,287$), loss to follow-up, or refusal. Second, cognitive measures fell into two categories—fluency and recall—and we excluded 2,294 participants who were missing all fluency information or all recall information at age 65. (Two-thirds of these were randomly selected, as part of the study protocol, to receive none of the neurocognitive testing.) Third, we excluded 1,258 participants who had valid scores for cognition at age 65, but had either left the study by age 72, or were missing all fluency or all recall information at age 72. Therefore, our analytic sample included 3,713 participants who had valid data on key variables at age 65 and age 72.

The analytic sample was more select than the original cohort. Participants in the analytic subsample were more likely to be women, had higher average standardized test scores in high school, and attained higher levels of education in adulthood than participants who were not in the analytic sample. There is less information about how cognitive performances in adulthood compared, because often excluded participants had not completed cognitive testing. However, where (incomplete) cognitive data did exist on participants who were not in the analytic sample, their scores were consistently and statistically significantly lower on all tests at both age 65 and 72.

Cognitive Functioning

Participants were asked to complete four cognitive tests at age 65 and 72: phonemic verbal fluency, semantic verbal fluency, and two tests of episodic memory. These tests were asked over the phone at age 65 and in-person at age 72. The phonemic verbal fluency, or letter fluency, test instructed respondents to think of as many words as they could starting with either "L" or "F" in 60 s; respondents did not receive credit for proper names, repeat words, or an already used word with a different ending (e.g., farmers, farming). The semantic verbal fluency, or category fluency, test also allotted each participant 60 s, requesting that they list as many words as fit into the category of either "foods" or "animals," depending on which category they were randomly assigned (Tombaugh et al., 1999). These two indicators of cognitive function are measured continuously, as number of words. A latent profile analysis (LPA) is the type of mixture model appropriate for such data (Collins & Lanza, 2010, table 1.1); however, LPA fails to detect the correct number of subgroups unless the indicator variables very clearly differentiate among the subgroups (Tein et al., 2013), which was not the case in our data (results available

upon request). Therefore, for both measures, we aggregated the continuous scores into quartiles of performance in order to be able to estimate the more computationally-friendly LTA. We retained the 2004 thresholds between quartiles in the 2011 data, so as to scale all measures on the same metric despite longitudinal changes in the average performance of the sample (Moeller, 2015).

Participants were then asked to complete the immediate and delayed recall tasks as measures of episodic memory. The interviewer first read the respondent a list of 10 words and then requested him or her to list those words back. The interviewer gave the respondent a point for each word they remembered during the immediate recall test. The interviewer then asked the respondent to again list those words, approximately 10 min later, to obtain a score for the delayed recall test (Brandt et al., 1988). As with the verbal fluency measures, we aggregated the continuous scores into quartiles.

Analytic Strategy

We used the program MPlus to estimate a series of mixture models using data on verbal fluency and episodic memory performance when the participants were age 65 and 72. Our analyses accounted for missing data on the cognitive measures by using full information maximum likelihood (FIML) estimation. The FIML approach generates a likelihood function for each participant and estimates the model based on those predictions.

Our analysis followed a three-step method (Asparouhov & Muthén, 2014; Collins & Lanza, 2010). First, we estimated cross-sectional LCAs at age 65 and 72. LCAs identify subgroups, or classes, of cognitive performance. We used the sample size-adjusted Bayesian inference criterion (BIC) as a fit statistic to determine the number of classes that best represented the data at each wave (Nylund et al., 2007). The original formula for BIC adjusts the model log-likelihood for both sample size and number of parameters estimated, and sample size-adjusted BIC alters the original sample size adjustment in a way that, in simulation studies, has been shown to provide superior model fit information for mixture models (Nylund et al., 2007). A BIC difference between two models of 0–2 indicates weak evidence to prefer the model with the smaller BIC, 2–6 indicates positive evidence, 6–10 indicates strong evidence, and a difference of 10 and greater constitutes very strong evidence (Raftery, 1995).

Once the number of classes was established, each participant received a probability of belonging to each class based upon their observed scores on the cognitive measures. The subsequent presentation of LCA results is based on assigning each participant to a single, most likely class, based on their highest value among these posterior probabilities.

Second, we tested for measurement invariance between 2004 and 2011. That is, we assessed whether the classes that best fit the 2004 data were quantitatively identical to

the classes that best fit the 2011 data. Measurement invariance indicates that while individual participants may transition from class to class over time, the characteristics of the classes themselves do not change over time. Conversely, measurement variance indicates global developmental processes whereby as participants age, the classes that best described the population at an earlier measurement occasion in the life course differ from the classes that best describe a later measurement occasion. Both may be true: Some classes may remain stable while others change (Collins & Lanza, 2010). To test the measurement structure of the solution, one estimates a series of models successively constraining the item-response probabilities of classes in 2004 to be equal to the item-response probabilities of classes in 2011. Again, we used sample size-adjusted BIC to select the measurement structure that best represented the data.

Third, we estimated a LTA model to examine how individuals' probability of class membership changed over the 7-year period between age 65 and 72. We did not constrain transition probabilities, whereby all combinations of the age 65 and 72 classes were possible.

Results

Latent Class Analyses

The LCAs for each wave established the groupings of participants with similar cognitive performance. As displayed in the top panel of Table 1, sample size-adjusted BIC favored a five-class model over a four-class model by 2.5 units, corresponding to weak positive evidence. We ultimately selected the four-class model, although the five-class model is displayed in Supplementary Table 2A for comparison. We selected the four-class model because in sensitivity analyses (not shown) with different subsamples, BIC consistently indicated a four-class solution. (For example, an arm of the WLS collected participants' genetic data to calculate polygenic scores, including a score for cognitive performance, based on genome-wide association data. However, genetic data were not available for all participants.)

Table 2 displays the class prevalences and the item-response probabilities, with class assignment based upon highest posterior probabilities. (Further information about entropy and the distribution of the highest posterior probabilities for this and other models is presented in Supplementary Table 3A.) The first class, which we labeled *Overall Low Performance*, comprised 22% of the sample. This class contained participants who had less than 5% probability of a top-quartile score on any of the four tests. Their probability of scoring in the lowest quartile was over 50% for every test except letter fluency, for which it was 39%. We named the second class, 29% of the sample, *Overall High Performance*. Members of this class had over a 50% chance of scoring in the top quartile on the fluency tests and over a 75% chance of scoring in the top quartile on the episodic memory tests.

Table 1. Summary of Fit Statistics for Model Selection (Wisconsin Longitudinal Study, $N = 3,713$)

Number of latent classes	Number of parameters estimated	G^2	df	BIC	ℓ
2004, age 65					
2	25	546.7	230	33964.2	-16919.1
3	38	345.3	217	33828.3	-16818.4
4	51	237.1	204	33785.8	-16764.3
5	64	169.2	191	33783.3	-16730.3
6	77	147.3	178	33827.0	-16719.4
2011, age 72					
2	25	578.2	230	34299.0	-17086.5
3	38	380.6	217	34166.9	-16987.7
4	51	277.4	204	34129.3	-16936.1
5	64	235.0	191	34152.4	-16914.9
6	77	198.6	178	24181.6	-16896.7
Measurement invariance					
Number of constrained classes	Number of parameters estimated	G^2	df	BIC	ℓ
0	111	9806.5	65,169	66410.7	-32925.5
1	99	9924.3	65,185	66407.8	-32954.3
2	87	9987.8	65,197	66418.3	-32989.8
3	75	10348.5	65,212	66696.4	-33159.1
4	63	10674.5	65,209	67061.8	-33372.1

Table 2. Four Latent Class Model of Cognitive Performance at Age 65 (Wisconsin Longitudinal Study, 2004, $N = 3,713$)

Assigned label	Latent class			
	Overall low	Overall high	High fluency	High episodic memory
Latent class prevalences	.22	.29	.23	.26
Item-response probabilities				
Letter fluency				
Lowest 25%	.39	.03	.02	.31
25%–50%	.41	.12	.14	.37
50%–75%	.17	.27	.30	.29
Highest 25%	.03	.58	.54	.03
Category fluency				
Lowest 25%	.52	.07	.09	.29
25%–50%	.24	.12	.15	.27
50%–75%	.20	.29	.37	.33
Highest 25%	.04	.51	.38	.12
Immediate recall				
Lowest 25%	.55	.00	.19	.00
25%–50%	.34	.03	.37	.14
50%–75%	.11	.18	.32	.30
Highest 25%	.00	.79	.12	.57
Delayed recall				
Lowest 25%	.59	.00	.32	.06
25%–50%	.27	.03	.34	.10
50%–75%	.15	.14	.33	.22
Highest 25%	.00	.83	.01	.61

The third class, *High Fluency Performance*, comprised 23% of the sample. This was a group of individuals who had an 84% chance of scoring above the 50th percentile on letter fluency and a 75% chance of scoring about the 50th percentile on category fluency. Their immediate

and delayed recall scores were more moderate. Finally, we labeled the fourth class, which comprised 26% of the sample, *High Episodic Memory Performance*. These participants had low-to-moderate scores on fluency, but a 57% chance of scoring in the top quartile on immediate

recall and a 61% chance of scoring in the top quartile on delayed recall.

A four-class model was the best fit for the data at age 72, as displayed in the middle panel of Table 1. Table 3 displays the class prevalences and item-response probabilities. The *Overall Low Performance* class from age 65 occurred again at age 72, comprising 26% of the sample. There were three new classes. *Older Overall High Performance* was a class comprised of 22% of the sample, comprising people who had a high probability of scoring in the top two quartiles on all four tests. Their delayed recall performance was notably strong, with group members having an 83% chance of scoring in the top quartile. The *Older High Fluency Performance* class, 30% of the sample, included people who had a high probability of scoring in the top 50% on fluency, but had more moderate recall scores. Finally, the *Older High Episodic Memory Performance* class accounted for 22% of the sample whose scores were predominantly in the middle 50% of the distribution. Their best performance was delayed recall, with a 46% chance of scoring in the top quartile. Their poorest performance was in category fluency, with a 36% chance of scoring in the lowest quartile.

Measurement Invariance

The second step of the analysis was to test whether the four classes identified at age 72 were statistically similar to

their counterparts at age 65. The bottom panel of Table 1 shows fit statistics for models that applied different sets of equality constraints on the item-response probabilities. The best-fitting model constrained one class, *Overall Low Performance*, to be invariant over time. The statistical boundaries of the other three classes changed over time, reflecting some developmental decline in cognitive performance. Figure 1 provides a visual representation of the seven unique classes. Relative to their age 65 counterparts, the *Older* classes had generally lower probabilities of top-quartile scores, with markedly lower probabilities of a top-quartile score on immediate recall.

Latent Transition Analysis

Next, the latent transition analyses established how individual participants transitioned among the classes over the 7-year period. Table 4 presents the transition probabilities conditional on age 65 class assignment. Most participants appear along the diagonal of the cross-tabulation of age 65 classes by age 72 classes, indicating that most participants remained in their respective subgroup as they aged. Members of *Overall Low Performance* at age 65 had a 82% chance of remaining in that class at age 72. Members of *Overall High Performance* at age 65 had a 60% chance of moving to *Older Overall High Performance* by age 72; members of *High Fluency Performance* at 65, a 77% chance of moving to *Older High Fluency Performance*; and

Table 3. Four Latent Class Model of Cognitive Performance at Age 72 (Wisconsin Longitudinal Study, 2011, N = 3,713)

Assigned label	Latent class			
	Overall low	Older overall high	Older high fluency	Older high episodic memory
Latent class prevalences	.26	.22	.30	.22
Item-response probabilities				
Letter fluency				
Lowest 25%	.39	.03	.03	.29
25%–50%	.41	.20	.27	.56
50%–75%	.17	.24	.33	.15
Highest 25%	.03	.44	.37	.00
Category fluency				
Lowest 25%	.52	.10	.16	.36
25%–50%	.24	.13	.20	.26
50%–75%	.20	.45	.40	.32
Highest 25%	.04	.32	.23	.04
Immediate recall				
Lowest 25%	.55	.01	.23	.00
25%–50%	.34	.09	.38	.27
50%–75%	.11	.64	.38	.65
Highest 25%	.00	.27	.01	.08
Delayed recall				
Lowest 25%	.59	.01	.34	.06
25%–50%	.27	.03	.36	.15
50%–75%	.15	.13	.29	.33
Highest 25%	.00	.83	.01	.46

members of *High Episodic Memory* at 65, a 68% chance of moving to *Older High Episodic Memory*.

There were two transitions that had transition probabilities of 30% or higher. These transitions indicated change that was more pronounced than that of other members of their age 65 class. Those in the *Overall High Performance* class at age 65 had a 39% chance of moving to the *Older High Fluency Performance* class at age 72, indicating particular losses in episodic memory. Additionally, those in the *High Episodic Memory Performance* class at age 65 had a 30% chance of moving to the *Overall Low Performance* class at age 72. These participants also experienced losses in episodic memory.

Finally, there were two transitions that had probabilities of 18% and 17%, respectively. Both groups gained in episodic memory, the first group moving from *High Fluency Performance* at age 65 to *Older Overall High Performance*, and the second group moving from *Overall Low Performance* at 65 to *Older High Episodic Memory Performance* at age 72.

Discussion

This study used mixture modeling techniques to identify subgroups of cognitive aging among Wisconsin Longitudinal Study participants based on their verbal fluency and episodic memory performance at age 65 and age 72. We found four subgroups at age 65 and four subgroups

at age 72. However, with the exception of the *Overall Low Performance* subgroup, the item-response probabilities changed as the participants aged, indicating an overall decline in performance with advancing aging. Most individuals remained in the same subgroup over the 7-year period, which we believe indicates several varieties of normative cognitive aging. A subset of individuals changed subgroups in ways indicating declining episodic memory, which could be suggestive of higher level of risk for future cognitive pathology. We discuss the implications of these results for the continued use of population health data to advance research on cognitive aging.

The Potential Utility of Mixture Modeling in the Field of Cognitive Aging

An interindividual differences approach to cognitive aging suggests that normative cognitive aging is diverse, with multiple distinct trajectories possible (Casaletto et al., 2019). A person-centered approach such as mixture modeling is an ideal way to explore this possibility (B. Muthén & L. K. Muthén, 2000). We found individual differences in both dispersion and in intraindividual change in the WLS data. The LCAs demonstrated four common patterns of cognitive aging in the domains of verbal fluency and episodic memory at age 65. With regard to dispersion, while there were participants who scored high on all four tests, as well as participants who scored low on all four tests, there were also participants who scored high on verbal fluency but not episodic memory, and vice versa. Prior research has identified dispersion as an independent risk factor for ADRD (Costa et al., 2019; Roalf et al., 2016). The WLS, however, lacks clinical measures of dementia or biomarker measures of neuropathology (although clinical measures are forthcoming [University of Wisconsin Center for Demography of Health and Aging, 2018]).

Nevertheless, this study’s findings with respect to dispersion may have some implications for the study of ADRD. By one estimate, in 2011 when these participants were age 72, 3% of Americans in their age group (65–74) had Alzheimer’s disease (Hebert et al., 2013). WLS participants were very unlikely to have had serious cognitive impairment; all participants were White and had obtained a high school degree or more, both of which are associated with lower rates of ADRD in the general population (Garcia

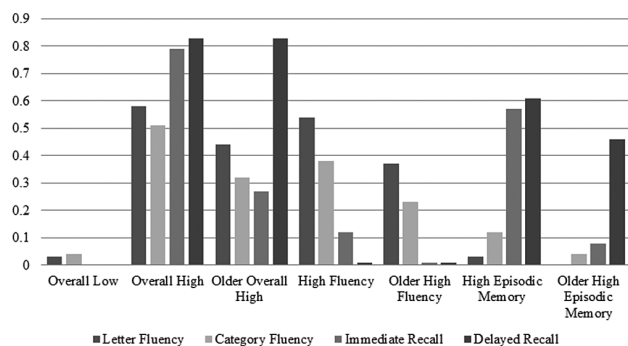


Figure 1. Probability of top-quartile cognitive test scores conditional on latent class membership (Wisconsin Longitudinal Study, N = 3,713). Note: Absent bars indicate a 0% probability of scoring in the top quartile of the respective test.

Table 4. Probability of Transitioning to Age 72 Class Conditional on Age 65 Class

	Age 72 class			
	Overall low	Older overall high	Older high fluency	Older high episodic memory
Age 65 class				
Overall low	.82	.00	.00	.17
Overall high	.00	.60	.39	.01
High fluency	.03	.18	.77	.02
High episodic memory	.30	.00	.02	.68

et al., 2020). Moreover, the lengthy interview would pose significant problems to a person with cognitive limitations. However, as the WLS participants advance into later life, their risk for ADRD increases considerably. Demographers estimate that among Americans born in 1940 and without dementia at age 70, 30.8% of men and 37.4% of women will develop it before they die (Fishman, 2017). Scenarios in which dementia could be delayed by 5 years in healthy 70-year-olds reduce those numbers considerably (Fishman, 2017).

Thus, an important question is whether mixture modeling techniques might help to identify people who are at risk of developing ADRD in the future within population health and epidemiological research. Declines in cognitive performance begin early among persons who later develop ADRD; in one study, decline in episodic memory began an average of 7 years before dementia diagnosis (Hamel et al., 2015). Previous mixture modeling studies have identified overall poor cognitive performance, as well as specific memory impairment, as risk factors for later dementia diagnosis and, upon autopsy, evidence of Alzheimer's pathology (Zammit, Hall, et al., 2019; Zammit, Muniz-Terrera, et al., 2019).

In addition to dispersion, there was diversity in intraindividual change. The WLS assessed cognition at age 65 and again at age 72. Our findings indicated that approximately three-quarters of people stayed in the same qualitative subgroup, but with the exception of the *Overall Low Performance* subgroup, each subgroup demonstrated quantitative decline. This result is consistent with the individual differences approach, which posits multiple trajectories within the "normal" range of cognitive aging (Hertzog, 2008). Other mixture modeling work further demonstrates that the rate of dementia diagnoses is highest among people who transition into a qualitatively distinct subgroup (Zammit et al., 2020).

Thus, we posit that our findings might illuminate such risk among participants who showed decline that was more pronounced than that of other members of their age 65 class. Specifically, those in the *High Episodic Memory Performance* class at age 65 had a 68% chance of becoming *Older High Episodic Memory Performance* at age 72, but also a 30% chance of moving to the *Overall Low Performance* class. Additionally, those in the *Overall High Performance* class at age 65 had a 39% chance of moving to the *Older High Fluency Performance* class at age 72. These transitions represent marked intraindividual decline in episodic memory.

Finally, we comment briefly on the two subgroups showing gains in episodic memory, the first group moving from *High Fluency Performance* at age 65 to *Older Overall High Performance*, and the second group moving from *Overall Low Performance* at 65 to *Older High Episodic Memory Performance* at age 72. As true improvement in episodic memory is unusual at this stage of life (Salthouse, 2019), these participants may have had performances on

the recall trials at age 65 that did not accurately represent their ability. Alternatively, improvements may be methodological artifacts, representing the effects of prior test experience (Salthouse, 2019).

Age and Cohort Effects on Population Subgroups of Cognition

Prior studies of individual differences in cognitive aging that have used mixture modeling with large social surveys have not focused on participant age, likely because it is unusual to have a sample from a single birth cohort. For example, the average participant in the Rush Memory and Aging Project was born in 1939, but the sample's standard deviation on age was over 7 years (Bennett et al., 2018; Zammit et al., 2020). Participants in the subsample of the HRS that Huang and colleagues analyzed were 78 or older in 2010, with 17% of the sample aged 89 or older (Huang et al., 2019). Both research groups found statistically invariant subgroups over time. In the WLS, which assesses developmental change within the birth cohort, the characteristics of subgroups did change over time. This finding indicates the importance of future studies seeking to document patterns of cognitive aging to attend to the ways in which age effects, or developmental changes, affect subgroups over time.

Cohort effects are a related issue: The sociohistorical contexts in which neurological development occurs may influence the subgroups of cognitive ability that emerge in adults in the population. In young adulthood and midlife, the Baby Boomer cohorts in the United States showed markedly stronger fluid intelligence skills than their parents had at similar ages, due to increased levels of educational attainment made possible by mandatory schooling, the G.I Bill, and other U.S. investments in education, alongside increasing work complexity for the population at large (Willis & Schaie, 2014). Moreover, sociohistorical contexts might also influence individual differences in cognitive decline and impairment. The Einstein Aging Study indicated a decline in the rate of dementia cases among participants in succeeding cohorts born before 1920, 1920–1925, 1925–1929, and later than 1929 (Derby et al., 2017), and analysis of a larger body of data concluded that rates of dementia are declining by birth cohort nationally (Leggett et al., 2019). With both age effects and cohort effects at play in cognitive aging, we might expect to identify different subgroups in different samples, rather than a universal set of subgroups.

Limitations

It is important to note several limitations of this study. First, the sample is limited to white older adults, all of whom graduated from high school. Given that educational

attainment has a strong association with cognitive aging (Seblova et al., 2020), this sample is unsuitable to describe subgroups that might be characteristic of older adults with less than a high school diploma. Moreover, rates of cognitive impairment are much higher in older adults from historically marginalized racial/ethnic groups than they are in white older adults, and older adults of color are more likely to be exposed to particular risk and protective factors throughout their life courses, such as racial discrimination (Mayeda et al., 2016). Accordingly, the subgroups identified in this analysis might not be applicable to samples of older adults from non-white racial/ethnic groups.

Second, similar to other methods of assessing intraindividual change, LTA models do not include a straightforward means of evaluating selective attrition because they require participation at all waves. The concern is that data might be missing not at random (MNAR) because individuals who dropped from the study and participants with missing data had poorer cognition than participants whose data are included in our subsample. We have descriptive evidence to address this concern. Among the 3,052 participants who left the study before age 65, we have no information on adult cognition, but their cognitive performance in adolescence was significantly lower than the adolescent performance of participants who were still involved with the study at age 65. Among participants who continued in the study until late midlife, but did not have valid cognitive scores at both age 65 and age 72, the cognition data that are available demonstrate significantly poorer performance than what was documented for the analytic subsample. However, within our work in the WLS using other statistical methods with less missing data stringency and more opportunities for assessing selective attrition, we have found little evidence that selective attrition biases results (Moorman et al., 2018, 2019). It remains possible, however, that selective attrition has biased estimates of the number of participants who transitioned from relatively more favorable subgroups to less favorable subgroups between ages 65 and 72.

Third, this study is limited to documenting patterns of cognitive aging in episodic memory and in verbal fluency. The Wisconsin Longitudinal Study lacks data on an exhaustive set of cognitive domains. Future research with data on additional domains, such as spatial ability, may yield additional or more complex subgroups of cognitive aging that reflect involvement of distinct parts of the brain.

Conclusion

Findings from our study contribute to a growing body of empirical evidence that mixture modeling is a feasible approach to assess later-life cognition, both at a single point in later life and throughout later life. The four subgroups identified among WLS participants, both at ages 65 and 72, indicated discrete categories of performance across multiple cognitive functions with interpretable quantitative

and qualitative differences. As large population surveys include various protocols for assessing cognition and ADRD (e.g., Langa et al., 2020), additional analyses that directly compare findings from mixture modeling to findings from clinical assessments will be possible. These analyses will help to further elucidate the utility of mixture modeling to advance the rapidly growing and increasingly multidisciplinary field of cognitive aging and ADRD research.

Supplementary Material

Supplementary data are available at *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences* online.

Funding

This work was supported by the National Institutes of Health (NIA R01 AG 057491).

Conflict of Interest

None declared.

Acknowledgments

This study was not preregistered. Wisconsin Longitudinal Study data are publicly available to researchers at <https://www.ssc.wisc.edu/wlsresearch/data/>. The Stata and MPlus command files necessary to replicate these analyses are available at openICPSR.org.

References

- Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using MPlus. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(3), 329–341. doi:10.1080/10705511.2014.915181
- Bennett, D. A., Buchman, A. S., Boyle, P. A., Barnes, L. L., Wilson, R. S., & Schneider, J. A. (2018). Religious Orders Study and Rush Memory and Aging Project. *Journal of Alzheimer's Disease*, 64(Suppl. 1), S161–S189. doi:10.3233/JAD-179939
- Brandt, J., Spencer, M., & Folstein, M. (1988). Telephone interview for cognitive status. *Neuropsychiatry, Neuropsychology, Behavioral Neurology*, 1, 111–117.
- Casaleto, K. B., Elahi, F. M., Staffaroni, A. M., Walters, S., Contreras, W. R., Wolf, A., Dubal, D., Miller, B., Yaffe, K., & Kramer, J. H. (2019). Cognitive aging is not created equally: Differentiating unique cognitive phenotypes in “normal” adults. *Neurobiology of Aging*, 77, 13–19. doi:10.1016/j.neurobiolaging.2019.01.007
- Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences* (Vol. 718). John Wiley & Sons.
- Costa, A. S., Dogan, I., Schulz, J. B., & Reetz, K. (2019). Going beyond the mean: Intraindividual variability of cognitive

- performance in prodromal and early neurodegenerative disorders. *The Clinical Neuropsychologist*, 33(2), 369–389. doi:10.1080/13854046.2018.1533587
- Costa, P. S., Santos, N. C., Cunha, P., Palha, J. A., & Sousa, N. (2013). The use of Bayesian latent class cluster models to classify patterns of cognitive performance in healthy ageing. *PLoS ONE*, 8(8), e71940. doi:10.1371/journal.pone.0071940
- Derby, C. A., Katz, M. J., Lipton, R. B., & Hall, C. B. (2017). Trends in dementia incidence in a birth cohort analysis of the Einstein Aging Study. *JAMA Neurology*, 74(11), 1345–1351. doi:10.1001/jamaneurol.2017.1964
- Fishman, E. (2017). Risk of developing dementia at older ages in the United States. *Demography*, 54(5), 1897–1919. doi:10.1007/s13524-017-0598-7
- Garcia, M. A., Downer, B., Chiu, C.-T., Saenz, J. L., Ortiz, K., & Wong, R. (2020). Educational benefits and cognitive health life expectancies: Racial/ethnic, nativity, and gender disparities. *The Gerontologist*. doi:10.1093/geront/gnaa112
- Ghisletta, P., Mason, F., von Oertzen, T., Hertzog, C., Nilsson, L.-G., & Lindenberger, U. (2020). On the use of growth models to study normal cognitive aging. *International Journal of Behavioral Development*, 44(1), 88–96. doi:10.1177/0165025419851576
- Goh, J. O., An, Y., & Resnick, S. M. (2012). Differential trajectories of age-related changes in components of executive and memory processes. *Psychology and Aging*, 27(3), 707–719. doi:10.1037/a0026715
- Gorus, E., De Raedt, R., & Mets, T. (2006). Diversity, dispersion and inconsistency of reaction time measures: Effects of age and task complexity. *Aging Clinical and Experimental Research*, 18(5), 407–417. doi:10.1007/BF03324837
- Hamel, R., Köhler, S., Sistermans, N., Koene, T., Pijnenburg, Y., van der Flier, W., Scheltens, P., Aalten, P., Verhey, F., Visser, P. J., & Ramakers, I. (2015). The trajectory of cognitive decline in the pre-dementia phase in memory clinic visitors: findings from the 4C-MCI study. *Psychological Medicine*, 45(7), 1509–1519. doi:10.1017/S0033291714002645
- Hebert, L. E., Weuve, J., Scherr, P. A., & Evans, D. A. (2013). Alzheimer disease in the United States (2010–2050) estimated using the 2010 census. *Neurology*, 80(19), 1778–1783. doi:10.1212/WNL.0b013e31828726f5
- Herd, P., Carr, D., & Roan, C. (2014). Cohort profile: Wisconsin Longitudinal Study (WLS). *International Journal of Epidemiology*, 43(1), 34–41. doi:10.1093/ije/dys194
- Hertzog, C. (2008). Theoretical approaches to the study of cognitive aging: An individual-differences perspective. In S. M. Hofer & D. F. Alwin (Eds.), *Handbook of cognitive aging: Interdisciplinary perspectives* (pp. 34–49). SAGE.
- Huang, F., Zhang, M., & Wang, S. (2019). Changes in cognitive function among older adults: A latent profile transition analysis. *Archives of Gerontology and Geriatrics*, 80, 12–19. doi:10.1016/j.archger.2018.09.006
- Hughes, M. L., Agrigoroaei, S., Jeon, M., Bruzzese, M., & Lachman, M. E. (2018). Change in cognitive performance from midlife into old age: Findings from the Midlife in the United States (MIDUS) study. *Journal of the International Neuropsychological Society*, 24(8), 805–820. doi:10.1017/S1355617718000425
- Karlamangla, A. S., Miller-Martinez, D., Aneshensel, C. S., Seeman, T. E., Wight, R. G., & Chodosh, J. (2009). Trajectories of cognitive function in late life in the United States: Demographic and socioeconomic predictors. *American Journal of Epidemiology*, 170(3), 331–342. doi:10.1093/aje/kwp154
- Langa, K. M., Ryan, L. H., McCammon, R. J., Jones, R. N., Manly, J. J., Levine, D. A., Sonnega, A., Farron, M., & Weir, D. R. (2020). The Health and Retirement Study harmonized cognitive assessment protocol project: Study design and methods. *Neuroepidemiology*, 54(1), 64–74. doi:10.1159/000503004
- Leggett, A., Clarke, P., Zivin, K., McCammon, R. J., Elliott, M. R., & Langa, K. M. (2019). Recent improvements in cognitive functioning among older U.S. adults: How much does increasing educational attainment explain? *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 74(3), 536–545. doi:10.1093/geronb/gbw210
- Lindenberger, U. (2014). Human cognitive aging: Corriger la fortune? *Science (New York, N.Y.)*, 346(6209), 572–578. doi:10.1126/science.1254403
- MacDonald, S. W. S., Li, S.-C., & Bäckman, L. (2009). Neural underpinnings of within-person variability in cognitive functioning. *Psychology and Aging*, 24(4), 792–808. doi:10.1037/a0017798
- Mayeda, E. R., Glymour, M. M., Quesenberry, C. P., & Whitmer, R. A. (2016). Inequalities in dementia incidence between six racial and ethnic groups over 14 years. *Alzheimer's & Dementia*, 12(3), 216–224. doi:10.1016/j.jalz.2015.12.007
- Mella, N., Fagot, D., & de Ribaupierre, A. (2016). Dispersion in cognitive functioning: Age differences over the lifespan. *Journal of Clinical and Experimental Neuropsychology*, 38(1), 111–126. doi:10.1080/13803395.2015.1089979
- Moeller, J. (2015). A word on standardization in longitudinal studies: don't. *Frontiers in Psychology*, 6, 1389. doi:10.3389/fpsyg.2015.01389
- Moorman, S. M., Carr, K., & Greenfield, E. A. (2018). Childhood socioeconomic status and genetic risk for poorer cognition in later life. *Social Science & Medicine* (1982), 212, 219–226. doi:10.1016/j.socscimed.2018.07.025
- Moorman, S. M., Greenfield, E. A., & Garcia, S. (2019). School context in adolescence and cognitive functioning 50 years later. *Journal of Health and Social Behavior*, 60(4), 493–508. doi:10.1177/0022146519887354
- Muthén, B., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism, Clinical and Experimental Research*, 24(6), 882–891. doi:10.1111/j.1530-0277.2000.tb02070.x
- Nesselroade, J. R. (1991). The warp and the woof of the developmental fabric. In L. S. Liben, R. M. Downs, & D. S. Palermo (Eds.), *Visions of aesthetics, the environment & development: The legacy of Joachim F. Wohlwill* (pp. 213–240). Lawrence Erlbaum Associates, Inc.
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, 14(4), 535–569. doi:10.1080/10705510701575396
- Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25, 111–163. doi:10.2307/271063
- Roalf, D. R., Quarmley, M., Mechanic-Hamilton, D., Wolk, D. A., Arnold, S. E., & Moberg, P. J.; Alzheimer's Disease Neuroimaging Initiative. (2016). Within-individual variability: An index for subtle change in neurocognition in mild cognitive impairment. *Journal of Alzheimer's Disease*, 54(1), 325–335. doi:10.3233/JAD-160259

- Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society*, 16(5), 754–760. doi:10.1017/S1355617710000706
- Salthouse, T. A. (2017). Contributions of the individual differences approach to cognitive aging. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 72(1), 7–15. doi:10.1093/geronb/gbw069
- Salthouse, T. A. (2019). Trajectories of normal cognitive aging. *Psychology and Aging*, 34(1), 17–24. doi:10.1037/pag0000288
- Seblova, D., Berggren, R., & Lövdén, M. (2020). Education and age-related decline in cognitive performance: Systematic review and meta-analysis of longitudinal cohort studies. *Ageing Research Reviews*, 58, 101005. doi:10.1016/j.arr.2019.101005
- Sharifian, N., Spivey, B. N., Zaheed, A. B., & Zahodne, L. B. (2020). Psychological distress links perceived neighborhood characteristics to longitudinal trajectories of cognitive health in older adulthood. *Social Science & Medicine (1982)*, 258, 113125. doi:10.1016/j.socscimed.2020.113125
- Tein, J. Y., Coxé, S., & Cham, H. (2013). Statistical power to detect the correct number of classes in latent profile analysis. *Structural Equation Modeling*, 20(4), 640–657. doi:10.1080/10705511.2013.824781
- Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology*, 14(2), 167–177. doi:10.1016/S0887-6177(97)00095-4
- University of Wisconsin Center for Demography of Health and Aging. (2018). *New grant for the Wisconsin Longitudinal Study*. <https://cdha.wisc.edu/2018/11/29/new-grant-for-the-wisconsin-longitudinal-study/>
- Willis, S. L., & Schaie, K. W. (2014). Cognitive functioning in the baby boomers: Longitudinal and cohort effects. In S. K. Whitbourne & S. L. Willis (Eds.), *The Baby Boomers grow up: Contemporary perspectives on midlife* (pp. 205–236). Routledge.
- Wisconsin Legislative Reference Bureau. (2017). 2015–2016 Wisconsin Blue Book. <http://legis.wisconsin.gov/LRB/publications/wisconsin-blue-book-2015/>
- Zammit, A. R., Bennett, D. A., Hall, C. B., Lipton, R. B., Katz, M. J., & Muniz-Terrera, G. (2020). A latent transition analysis model to assess change in cognitive states over three occasions: Results from the rush memory and aging project. *Journal of Alzheimer's Disease*, 73(3), 1063–1073. doi:10.3233/JAD-190778
- Zammit, A. R., Hall, C. B., Bennett, D. A., Ezzati, A., Katz, M. J., Muniz-Terrera, G., & Lipton, R. B. (2019). Neuropsychological latent classes at enrollment and postmortem neuropathology. *Alzheimer's & Dementia*, 15(9), 1195–1207. doi:10.1016/j.jalz.2019.05.012
- Zammit, A. R., Hall, C. B., Katz, M. J., Muniz-Terrera, G., Ezzati, A., Bennett, D. A., & Lipton, R. B. (2018). Class-specific incidence of all-cause dementia and Alzheimer's disease: A latent class approach. *Journal of Alzheimer's Disease*, 66(1), 347–357. doi:10.3233/JAD-180604
- Zammit, A. R., Hall, C. B., Lipton, R. B., Katz, M. J., & Muniz-Terrera, G. (2018). Identification of heterogeneous cognitive subgroups in community-dwelling older adults: A latent class analysis of the Einstein Aging Study. *Journal of the International Neuropsychological Society*, 24(5), 511–523. doi:10.1017/S135561771700128X
- Zammit, A. R., Muniz-Terrera, G., Katz, M. J., Hall, C. B., Ezzati, A., Bennett, D. A., & Lipton, R. B. (2019). Subtypes based on neuropsychological performance predict incident dementia: Findings from the Rush Memory and Aging Project. *Journal of Alzheimer's Disease*, 67(1), 125–135. doi:10.3233/JAD-180737