



Published in final edited form as:

Exp Aging Res. 2013 ; 39(5): . doi:10.1080/0361073X.2013.839029.

QUALITATIVE NEUROPSYCHOLOGICAL MEASURES: NORMATIVE DATA ON EXECUTIVE FUNCTIONING TESTS FROM THE FRAMINGHAM OFFSPRING STUDY

Lisa D. Hankee

Boston University School of Medicine, Boston, Massachusetts, USA Framingham Heart Study/
National Heart Lung and Blood Institute, Framingham, Massachusetts, USA

Sarah R. Preis

Framingham Heart Study/National Heart Lung and Blood Institute, Framingham, Massachusetts,
USA Boston University School of Public Health, Boston, Massachusetts, USA

Alexa S. Beiser

Boston University School of Medicine, Boston, Massachusetts, USA Framingham Heart Study/
National Heart Lung and Blood Institute, Framingham, Massachusetts, USA Boston University
School of Public Health, Boston, Massachusetts, USA

Sherral A. Devine, Yulin Liu, Sudha Seshadri, Philip A. Wolf, and Rhoda Au

Boston University School of Medicine, Boston, Massachusetts, USA Framingham Heart Study/
National Heart Lung and Blood Institute, Framingham, Massachusetts, USA

Abstract

Background/Study Context—Studies have found that executive functioning is affected early in the pathophysiological processes associated with Alzheimer’s disease and vascular dementia. There also exists a range of functioning on executive tasks during normal aging. Although qualitative data are commonly utilized in clinical practice for evaluating subtle changes in cognitive functioning and diagnostic discernment, it is not clear whether error responses used in clinical practice are also evident as normative behavior.

Methods—As part of an extensive battery of neuropsychological tests, executive functioning measures (i.e., Trail Making-B, Similarities and Verbal Fluency tests) were administered via standardized administration prescript. Regression analyses were used to determine associations between vascular aging indices and qualitative performance measures. Descriptive statistics are included for 1907 cognitively normal individuals.

Results—Results suggest that while qualitative errors do occur, they are relatively infrequent within a presumably cognitively normal sample. Error commission rates on executive functioning tests are significantly associated with both age and education.

Conclusion—Provided is a baseline profile of errors committed on tests of executive function across a range of age and educational levels. The normative datasets are included, stratified by age and educational achievement, for which to compare qualitative test performance of clinical and research populations.

Introduction

Changes in neurological structure and function are associated with the aging process (Fjell & Walhovd, 2010). Research on volumetric changes associated with aging have reported age-related declines in total cerebral brain volume (Peters, 2006; Walhovd et al., 2011), while emerging evidence shows a tendency for earlier and preferential disruption of prefrontal systems due to normal aging (Braver et al., 2001, McDaniel & Einstein, 2011, Raz et al., 1997) in concordance with the “frontal aging hypothesis” (West, 1996). Accelerated rates of decline have been noted in the frontal cortex and prefrontal areas as compared to the temporal, parietal, and occipital cortices (Drag & Bieliauskas, 2010). Further, compounding vascular risk factors are evident as early as middle age (Gouw et al., 2008; Raz et al., 2012; Wen et al., 2009), with the brains of almost all healthy adults showing at least some white matter hyperintensities (WMH, Sachdev et al., 2008) and an accelerated rate of expansion of frontal WMH which exceeds that observed within other lobes (Gouw et al., 2008).

In concert with these neurological changes, there are also age-related changes evident on cognitive test performance. Through the normal aging process, declines are evident in information processing efficiency (Salthouse, 1996), working memory (Hasher et al., 2007) and various tasks requiring executive functions (e.g., inhibitory control [Chao & Knight, 1997]; planning [Sanders & Schmitter-Edgecombe, 2012]), while cognitive heterogeneity has been reported on tasks of executive functioning, attention, and select non-verbal abilities (Ardila, 2007).

Prior studies from the Framingham Heart Study (FHS) show significant and positive associations between total cerebral brain volume (TCBV) and performance on tests of attention and executive function (i.e., Trail-making test; Seshadri et al., 2004). Vascular risk factors in particular are associated with heightened susceptibility to performance decline on formal measures of executive functioning. Recent analyses by FHS linked deficits on tests of executive functions to the presence of vascular risk factors (i.e., based on the Framingham Stroke Risk Profile [FSRP, Wolf et al., 1991]). Midlife hypertension has been associated with accelerated progression of WMH and worsening executive functioning (DeBette et al., 2011). Tan and colleagues (2011) found that diabetes, elevated glycohemoglobin, HOMA-IR, and fasting insulin were related to poorer executive function scores.

As the population ages, an additional concern is the heightened potential for neurodegenerative diseases (Hebert et al., 1995; Kukull, et al., 2002), particularly Alzheimer’s disease (AD). Recent evidence suggests subtle performance decrement at least a decade prior to clinically meaningful cognitive changes (Chen et al., 2000; Howieson et al., 2008). These findings have resulted in the expectation that disease-modifying treatments will have optimal effects when administered prior to significant cognitive impairment (i.e., at a “presymptomatic” or “preclinical” stage, Sperling et al., 2011), and calls have been made for the development of highly sensitive tests to facilitate early detection and delineation of factors that may predict the impending subsequent cognitive and functional decline.

In response to the increasing interest in pre-clinical cognitive changes, FHS implemented a qualitative coding system to be used in conjunction with the standard quantitative scores for a battery of neuropsychological tests, including several commonly used tests of executive functioning. Qualitative analyses of executive functioning tasks may provide subtle indications of early neuropathological changes. Research has shown that deficits in executive function are a useful predictor in determining cognitively healthy individuals at risk for developing dementia (Ritchie et al., 2001), and tests of executive function and

attention best discriminated between non-converters and incident AD cases after a 4-year time interval (Rapp & Reischies, 2005). In addition, studies of cognitively-healthy individuals with genetic risk for AD (i.e., APOE4 positive) have demonstrated a heightened vulnerability for qualitative error commission on executive function tests of inhibition and cognitive flexibility (Wetter et al., 2005) and an asymmetric performance on verbal fluency tasks (Houston et al., 2005). Qualitative observations are known to be clinically-rich and often particularly helpful in diagnostic considerations and treatment recommendations (Kaplan, 1988; Lezak, 2004).

The Boston Process Approach (Kaplan, 1988) is a method for analyzing NP test performance, which, taken in concert with more traditional test performance data, adds sensitivity and meaning to neuropsychological assessment (Milberg & Hebben, 2006; Poreh, 2006). Despite the use of qualitative scoring in clinical neuropsychology practice, there is paucity of application in research settings, and there have not been any large-scale epidemiologic studies that have integrated both quantitative and qualitative scoring techniques as part of their standardized neuropsychological testing protocol. Rather, qualitative assessment of test performance to date has relied on subjective clinical judgment as opposed to more objective comparison of error scores to a normative sample. Analysis of the qualitative aspects of an individual's performance, with systematic coding and scoring, and in conjunction with traditional quantitative measures facilitates the examination of error commission as normative behaviors. The focus of this study is two-fold: first, to provide normative data for qualitative error measures on several commonly used executive function tests, and second, to relate these qualitative measures to vascular aging indices (i.e., total and frontal brain volume, white matter hyperintensities volume, silent cerebral infarcts).

Methods

Participants

Established in 1948, the Framingham Heart Study recruited an Original cohort for biennial examination to identify risk factors of cardiovascular disease. In 1971, biological Offspring of the Original FHS Cohort and their spouses were invited to regular health examinations approximately every four years (Kannel et al., 1979). From 2005-2008, 1,993 Offspring participants took part in a follow-up study on cognition. Participants with prevalent clinical stroke, dementia, or other neurological diseases (e.g., severe head trauma, multiple sclerosis, etc) were excluded (n=86). A total of 1907 participants (54% women) comprised the sample for this normative study.

The Institutional Review Board (IRB) at Boston University Medical Center (BUMC) approved the study protocol. Informed consent was obtained from all participants. Table 1 provides demographic information on the study sample.

Magnetic Resonance Imaging Measures

The brain MRI acquisition and assessment techniques utilized in the Framingham Heart Study offspring cohort have been described in detail elsewhere (Seshadri et al., 2004). Briefly, MR images were taken with a 1-T field strength Siemens Magnetom scanner, using a double spin-echo coronal imaging sequence to acquire 4-mm contiguous slices. Frontal lobar volume, total cerebral brain volume, large white matter hyperintensities (WMH-L) volume and the presence of silent cerebral infarcts (SCI) were computed using methods that have been previously described and validated (DeCarli et al., 2005).

Neuropsychological Executive Functioning Measures

Participants completed four tests of executive functioning as part of a larger neuropsychological test battery (see Au et al., 2004). The Trailmaking Test-part B (TrB; Army Individual Test Battery, 1944), category fluency (Animal naming) and phonemic fluency (FAS) from the Controlled Oral Word Association Test (COWAT; Benton et al., 1994), and the Wechsler Adult Intelligence Scale Similarities subtest (Wechsler, 1955) were administered by examiners trained in standardized administration protocol. Normative qualitative data tables for these four widely used executive functioning measures are included. Qualitative norms for the remainder of the tests administered are available in an on-line supplement.

Qualitative Scoring Protocol

Table 2 shows the qualitative error types for each executive function test.

A combination of perceptual, sequencing and set-shifting errors were included in the Trailmaking Test-part B (TrB) total qualitative errors. Errors associated with deficits in executive planning skills (i.e., pen lifts) were also coded for the TrB test.

On the tests of verbal fluency (*FAS*, *Animals*), error categories included perseverative errors and set-maintenance errors. For each trial of *FAS*, words that were repeated at any time during the trial (i.e., did not have to be repeated sequentially) were qualitatively scored as perseverations while the number of broken rules (e.g., proper names, a word given that is the same as a previous response albeit with a different suffix) and words beginning with a letter other than F, A or S (for each respective trial) were summed for measures of errors due to loss of set. Similarly, for the *Animals* trial, repeated animals (i.e., sequential and non-sequential repetitions) were coded as perseverations and any word given that was not an animal was registered as a set-maintenance error. The total number of responses produced is also reported here, to serve a measure of verbal output and to enable percent accuracy classifications.

Responses on the WAIS Similarities subtest were qualitatively coded for concrete responses and set-maintenance errors. Concrete responses were coded when the participant apparently understood what was required but failed to give an abstract response (e.g., “food” for *egg-seed*). Set maintenance errors were coded when the participant either failed to explain (abstractly or concretely) how two items were alike or gave a response that explained how the two objects were related to each other (e.g., “a fly could land on a tree”). In each of the instances, the response produced is inconsistent with the task as directed.

To ensure accurate test administration and data collection, test sessions were digitally recorded and participant responses were transcribed verbatim. Monthly quality control (QC) reviews were conducted by a supervising neuropsychologist (SD) and post-doctoral fellow (LH). QC procedures involved listening to the digital voice recordings to check for accurate administration and transcription of verbal information. Additionally, each quantitative and qualitative variable was examined for scoring and data collection precision. QC checks were evenly distributed across all examiners and test batteries were randomly selected. QC feedback was communicated directly with the test examiner, and a QC metric (i.e., percent accuracy score) was calculated. Across all tests administered, the mean QC metric score was 98.9% (range of 90.9-99.9%), suggesting high inter-examiner consistency for these measures.

Statistical Analyses

For all variables, means and standard deviations were calculated for the entire study sample as well as by age group (<55, 55-64, 65-74, 75 years) and by education group (<High school diploma, High school diploma, College degree, Graduate degree). We used linear regression models to analyze the association between TCBV, frontal lobar volume, WMH-L volume, SCI and each of the executive functioning measures, adjusted for age and education. All statistical analyses were done using SAS version 9.2 (Cary, NC).

Results

Tables 3a-3b contain the total number of participants who completed the test, mean scores and standard deviations, stratified by age and educational attainment, for the *Trail Making Test (TrB)*. Qualitative errors appeared to vary significantly across age and education. Planning errors (i.e., pen lifts) were the most commonly observed errors (mean 1.4 ± 2.2). Qualitative measures were significantly related to age and educational achievement ($p < 0.0001$).

Tables 4a-4b contain the mean scores and standard deviations, stratified by age and educational attainment, for the *Verbal Fluency* test. Qualitative errors on phonemic fluency (FAS combined), when taken as a percentage of total verbal output, had associated age and education effects ($p < 0.0001$). FAS perseveration errors were also significantly associated with age ($p < 0.0001$) and education ($p = 0.006$). Qualitative errors on Animal naming, when considered in relation to the total number of animals produced, was significantly associated with age ($p < 0.001$) and education ($p = 0.007$). Animal perseverations also had significant relation to age ($p = 0.01$) and education effects ($p = 0.007$).

Tables 5a-5b contain the mean scores and standard deviations, stratified by age and educational attainment, for the *WAIS Similarities test*. Qualitative errors due to concrete thinking were observed most often (mean 4.4 ± 1.8) while set maintenance errors were observed less frequently (mean 0.7 ± 1.1). Error commission due to concrete responding and set loss were significantly associated with both age group and educational attainment ($p < 0.0001$).

Table 6 has information on the relationship between each test and MRI indices of aging. Total cerebral brain volume (TCBV) and frontal lobar volume (FLV) were significantly associated with several traditional quantitative measures. TCBV had a significant association with TrB time to completion ($p < 0.01$), FAS total responses ($p < 0.05$) and WAIS Similarities total score ($p < 0.01$), while FLV was significantly related to TrB time to completion ($p < 0.001$) and FAS total responses ($p < 0.05$). Significant associations were also found between FLV and qualitative measures on the TrB test. FLV was significantly associated with TrB total errors committed ($p < 0.01$) and with the log-transformed frequency of TrB pen lifts ($p < 0.01$). An association between WMH-L and TrB total errors was also evident ($p < 0.01$).

Discussion

Results demonstrate that error commission on neuropsychological tests of executive functioning is evident as normative behavior. Although occurrence is relatively infrequent, healthy aging participants make errors on executive functioning tests. Significant age and education effects were found for all test measures. The most commonly observed errors on executive functioning tests included perseveration (i.e., repetition) errors on a test of verbal fluency, pen lifts on the Trail-making Test-B, and errors due to concrete responding on WAIS Similarities subtest. Measures of total brain volume and frontal lobar volume were

significantly associated with all traditional quantitative measures for Trailmaking Test-B in addition to two qualitative indices, total errors committed and a measure of impulsivity (i.e., pen lifts). Also noted was an association between the volume of large white matter hyperintensities and TrB total errors.

It is important to note that the Offspring cohort is fairly well-educated. Approximately 40% of the sample had attained at least a college degree. The lower education range, especially those with less than a high school education, may be under-represented by this study sample. Also, as the FHS Offspring cohort is comprised of people of European descent, it does not adequately reflect the diverse ethnicities of the broader population, and therefore these normative data cannot be generalized to non-Caucasian populations. In addition, the participants' longstanding involvement with the study, including frequent physical health examinations, may result in a heightened awareness for cardiovascular and other health factors. Approximately 13% of the Offspring cohort has a history of cardiovascular disease, as compared to a prevalence of greater than one-third nationally (American Heart Association, 2012). Finally, the diagnostic value of qualitative neuropsychological data has yet to be determined, and the clinical significance of these qualitative measures will require longitudinal follow-up.

It bears mentioning that any sample of presumably cognitively-normal individuals may be contaminated with some individuals in the very early stages of a neurodegenerative disease process (Sliwinski et al., 1996; De Santi et al., 2008), especially for the older age group (i.e., 75+). When those experiencing pre-clinical levels of impairment or undetected cognitive decline are included in normative distribution, the result may be an overestimate of the actual population ranges for qualitative measures of error commission.

Despite these limitations, these analyses suggest that qualitative data can be accurately collected (i.e., scored and coded), and when considered in comparison to a cognitively healthy cohort, data derived from the qualitative analyses of neuropsychological test performance may provide meaningful information for identification of early, subtle signs of cognitive difficulty on executive function tests. Additional research is necessary to document formal test reliability for these measures. Finally, longitudinal follow-up is essential for determining if qualitative error commission is a preclinical marker in those who are subsequently identified with mild cognitive impairment (MCI) or who meet diagnostic criteria for dementia, and to enable analyses of the predictive validity these measures have for dementia and other neurodegenerative disorders.

Conclusion

This study provides a baseline profile of error commission rates across a range of age and educational levels. These data allow for normative comparison of the performance of clinical and research populations to evaluate qualitative performance on executive functioning tests.

Acknowledgments

This work was supported by the Framingham Heart Study's National Heart, Lung, and Blood Institute contract (N01-HC-25195), by grants (R01-AG16495, R01-AG08122, R01-AG033040) from the National Institute on Aging, and by grant (R01-NS17950) from the National Institute of Neurological Disorders and Stroke.

The authors thank the extraordinary participants and families of the Framingham Heart Study who made this work possible. We also acknowledge the great work of all the research assistants and study staff.

References

- American Heart Association. Heart disease and stroke statistics 2012 update: A report from the American Heart Association. *Circulation*. 2012; 125:e2–e220. [PubMed: 22179539]
- Army Individual Test Battery. Manual of Directions and Scoring. War Department, Adjutant General's Office; Washington, DC: 1944.
- Ardila A. Normal aging increases cognitive heterogeneity: Analysis of dispersion in WAIS-III scores across age. *Archives of Clinical Neuropsychology*. 2007; 22(8):1003–1011. [PubMed: 17904332]
- Au R, Seshadri S, Wolf PA, Elias MF, Elias PK, Sullivan L, Beiser A, D'Agostino R. New norms for a new generation: Cognitive performance in the Framingham Offspring Cohort. *Experimental Aging Research*. 2004; 30:333–358. [PubMed: 15371099]
- Benton, AL.; Hamsher, K.; Sivan, AB. Multilingual Aphasia Examination. 3rd ed. AJA Associates; Iowa City, IA: 1994.
- Braver TS, Barch DM, Keys BA, Carter CS, Cohen JD, Kaye JA, Janowsky JS, Taylor SF, Yesavage JA, Mummenthaler S, Jagust WJ, Reed BR. Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General*. 2001; 130:746–763. [PubMed: 11757878]
- Chao L, Knight R. Prefrontal deficits in attention and inhibitory control with aging. *Cerebral Cortex*. 1997; 7(1):63–69. [PubMed: 9023433]
- Chen P, Ratcliff G, Belle SH, Cauley JA, DeKosky ST, Ganguli M. Cognitive tests that best discriminate between presymptomatic AD and those who remain nondemented. *Neurology*. 2000; 55(12):1847–1853. [PubMed: 11134384]
- Debette S, Seshadri S, Beiser A, Au R, Himali JJ, Palumbo C, Wolf PA, DeCarli C. Midlife vascular risk factor exposure accelerates structural brain aging and cognitive decline. *Neurology*. 2011; 77:461–468. [PubMed: 21810696]
- DeCarli C, Massaro J, Harvey D, Hald J, Rullberg M, Au R, Beiser A, D'Agostino R, Wolf PA. Measures of brain morphology and infarction in the Framingham Heart Study: Establishing what is normal. *Neurobiology of Aging*. 2005; 26:491–510. [PubMed: 15653178]
- De Santi S, Pirraglia E, Barr W, Babb J, Williams S, Rogers K, Glodzik L, Brys M, Mosconi L, Reisberg B, Ferris S, de Leon MJ. Robust and conventional neuropsychological norms: Diagnosis and predictions of age-related cognitive decline. *Neuropsychology*. 2008; 22(4):469–484. [PubMed: 18590359]
- Drag LL, Bieliauskas LA. Contemporary review 2009: Cognitive aging. *Journal of Geriatric Psychiatry and Neurology*. 2010; 23:74–93.
- Fjell AM, Walhovd KB. Structural brain changes in aging: Courses, causes and cognitive consequences. *Reviews in the Neurosciences*. 2010; 21(3):187–221. [PubMed: 20879692]
- Gouw AA, van der Flier WM, Fazekas F, van Straaten EC, Pantoni L, Poggesi A, Inzitari D, Erkinjuntti T, Wahuund LO, Wakleamar G, Schmidt R, Scheltens P, Barkhof F. Progression of white matter hyperintensities and incidence of new lacunes over a 3-year period: The leukoaraiosis and disability study. *Stroke*. 2008; 39:1414–1420. [PubMed: 18323505]
- Hasher, L.; Lustin, C.; Zacks, RT. Inhibitory mechanisms and the control of attention. In: Conway, A.; Jarrold, C.; Kane, M.; Towse, J., editors. *Variation in Working Memory*. Oxford University Press; New York: 2007. p. 227-249.
- Hebert LE, Scherr PA, Beckett LA, Albert MS, Pilgrim DM, Chown MJ, Funkenstein HH, Evans DA. Age-specific incidence of Alzheimer's disease in a community population. *Journal of the American Medical Association*. 1995; 273:1354–1359. [PubMed: 7715060]
- Houston WS, Delis DC, Lansing A, Jacobson MW, Cobell KR, Salmon DP, Bondi MW. Executive function asymmetry in older adults genetically at-risk for Alzheimer's disease : Verbal fluency versus design fluency. *Journal of the International Neuropsychological Society*. 2005; 11(7):863–870. [PubMed: 16519265]
- Howieson DB, Carlson NE, Moore MM, Wasserman D, Abendroth CD, Payne-Murphy J, Kaye JA. Trajectory of mild cognitive impairment onset. *Journal of the International Neuropsychological Society*. 2008; 14:192–198. [PubMed: 18282317]

- Kannel WB, McGee DL. Diabetes and cardiovascular risk factors: The Framingham Study. *Circulation*. 1979; 59:8–13. [PubMed: 758126]
- Kaplan, E. A process approach to neuropsychological assessment. In: Boll, T.; Bryant, BK., editors. *Clinical neuropsychology and brain function: Research, measurement, and practice*. American Psychological Association; Washington D.C.: 1988.
- Kukull WA, Higdon R, Bowen JD, McCormick WC, Teri L, Schellenberg GD, van Belle G, Jolley L, Larson EB. Dementia and Alzheimer disease incidence: A prospective cohort study. *Archives of Neurology*. 2002; 59:1737–1746. [PubMed: 12433261]
- Lezak, M. *Neuropsychological Assessment*. Oxford University Press; New York: 2004.
- McDaniel MA, Einstein GO. The neuropsychology of prospective memory in normal aging: A componential approach. *Neuropsychologia*. 2011; 49(8):2147–2155. [PubMed: 21192957]
- Milberg, W.; Hebben, N. The historical antecedents of the Boston Process Approach. In: Poreh; Amir, M., editors. *The quantified process approach to neuropsychological assessment*. Taylor & Francis; Philadelphia, PA, US: 2006.
- Peters R. Aging and the brain. *Postgraduate Medical Journal*. 2006; 82:84–88. [PubMed: 16461469]
- Poreh, AM., editor. *The quantified process approach to neuropsychological assessment*. Taylor & Francis; Philadelphia, PA, US: 2006.
- Rapp MA, Reischies FM. Attention and executive control predict Alzheimer disease in late life: Results from the Berlin Aging Study (BASE). *American Journal of Geriatric Psychiatry*. 2005; 13(2):134–141. [PubMed: 15703322]
- Raz N, Gunning FM, Head D, Dupuis JH, McQuain J, Briggs SD, Loken WJ, Thornton AE, Acker JD. Selective aging of the human cerebral cortex observed in vivo: Differential vulnerability of the prefrontal gray matter. *Cerebral Cortex*. 1997; 7:268–282. [PubMed: 9143446]
- Raz N, Yang Y, Dahle CL, Land S. Volume of white matter hyperintensities in healthy adults: Contribution of age, vascular risk factors, and inflammation-related genetic variants. *Biochimica et Biophysica Acta*. 2012; 1822:361–369. [PubMed: 21889590]
- Ritchie K, Artero S, Touchon J. Classification criteria for mild cognitive impairment: A population-based validation study. *Neurology*. 2001; 56(1):37–42. [PubMed: 11148233]
- Sachdev P, Chen X, Wen W. White matter hyperintensities in mid-adult life. *Current Opinion in Psychiatry*. 2008; 21(3):268–274. [PubMed: 18382226]
- Sanders C, Schmitter-Edgecombe M. Identifying the nature of impairment in planning ability with normal aging. *Journal of Clinical and Experimental Neuropsychology*. 2012; 34(7):724–737. [PubMed: 22506736]
- Salthouse T. The processing-speed theory of adult age differences in cognition. *Psychological Review*. 1996; 103:403–428. [PubMed: 8759042]
- SAS Institute Inc.. *SAS® 9.2 Enhanced Logging Facilities*. SAS Institute Inc.; Cary, NC: 2008.
- Seshadri S, Wolf PA, Beiser A, Elias MF, Au R, Kase CS, D'Agostina RB, DeCarli C. Stroke risk profile, brain volume and cognitive function: The Framingham Offspring Study. *Neurology*. 2004; 63(9):1591–1599. [PubMed: 15534241]
- Sliwinski M, Lipton RB, Buschke H, Stewart W. The effects of preclinical dementia on estimates of normal cognitive functioning in aging. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences*. 1996; 51(4):217–225.
- Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR, Kaye J, Montine TJ, Park DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K, Carrillo MC, Thies B, Morrison-Bogorad M, Wagster MV, Phelps CH. Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dementia*. 2011; 7(3):280–292.
- Tan ZS, Beiser AS, Fox CS, Au R, Himali JJ, Debette S, DeCarli C, Vasan RS, Wolf PA, Seshadri S. Association of metabolic dysregulation with volumetric brain magnetic resonance imaging and cognitive markers of subclinical brain aging in middle-aged adults. *Diabetes Care*. 2011; 34:1766–1770. [PubMed: 21680719]

- Walhovd KB, Westlye LT, Amlien I, Espeseth T, Reinvang I, Raz N, Agartz I, Salat DH, Greve DN, Fischl B, Dale AM, Fjell AM. Consistent neuroanatomical age-related volume differences across multiple samples. *Neurobiology of Aging*. 2011; 32(5):916–932. [PubMed: 19570593]
- Wechsler, D. *Manual for the Wechsler Adult Intelligence Scale*. The Psychological Corporation; New York: 1955.
- Wen W, Sachdev PS, Li JJ, Chen X, Anstey KJ. White matter hyperintensities in the forties: Their prevalence and topography in an epidemiological sample aged 44-48. *Human Brain Mapping*. 2009; 30(4):1155–1167. [PubMed: 18465744]
- West RL. An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*. 1996; 120:272–292. [PubMed: 8831298]
- Wetter SR, Delis DC, Houston WS, Jacobson MW, Lansing A, Cobell K, Salmon DP, Bondi MW. Deficits in inhibition and flexibility are associated with the APOE- 4 allele in nondemented older adults. *Journal of Clinical and Experimental Neuropsychology*. 2005; 27(8):943–952. [PubMed: 16207619]
- Wolf PA, D’Agostino RB, Belanger AJ, Kannel WB. Probability of stroke: A risk profile from the Framingham study. *Stroke*. 1991; 22:312–318. [PubMed: 2003301]

Table 1

Study sample characteristics

	N=1907
<i>Categorical characteristics, n (%)</i>	
Age	
< 55	135 (7.1)
55-64	667 (35.0)
65-74	649 (34.0)
75	456 (23.9)
Sex	
Women	1030 (54.0)
Men	877 (46.0)
Educational Level	
< HS Diploma	64 (3.4)
HS Diploma	1079 (56.6)
College Degree	404 (21.2)
Graduate Degree(s)	360 (18.9)
Silent cerebral infarct (≥ 3 mm) *	202 (13.4)
History of CVD	246 (12.9)
Diabetes	244 (12.8)
Current smoking	148 (7.9)
<i>Continuous characteristics, mean (sd)</i>	
Total cerebral brain volume *	78.9 (3.9)
Frontal lobe volume *	36.1 (3.4)
Log WMH-L volume *	-2.4 (1.1)
Systolic blood pressure, mm Hg	128 (17)
Body mass index, kg/m ²	28.3 (5.3)
LDL cholesterol, mg/dL	106 (31)
HDL cholesterol, mg/dL	57 (18)
MMSE Score	28.7 (1.7)

Log WMH-L, volume large white matter hyperintensities, log-transformed.

* N=1503 for subset with MRI measures.

Table 2

Types of qualitative errors on executive functioning tests

Test	Error type	Range of errors
Trailmaking Test-B	Qualitative errors	0 – max
	Planning errors (pen lifts)	0 – max
Verbal Fluency: FAS	Perseveration errors	0 – max
	Set maintenance errors	0 – max
Verbal Fluency: Animal Naming	Perseveration errors	0 – max
	Set maintenance errors	0 – max
WAIS Similarities	Concrete responses	0 – 13
	Set maintenance errors	0 – 10

Note: WAIS, Wechsler Adult Intelligence Scale.

Table 3a

Trail Making Test-B qualitative data [mean (standard deviation)] stratified by age

Age (years)	<55	55-64	65-74	75	Total	F-value*	P-value
Completion time	66.7(25.0)	74.2(34.8)	95.0(48.5)	124.4(57.3)	91.8(49.1)	116.8	<0.0001
Total errors	0.4(0.7)	0.4(0.8)	0.7(1.0)	0.9(1.3)	0.6(1.0)	21.7	<0.0001
Log Pen lifts	0.7(0.9)	1.0(1.5)	1.5(2.3)	2.3(2.9)	1.4(2.2)	28.2	<0.0001
	<i>n=130</i>	<i>n=652</i>	<i>n=616</i>	<i>n=393</i>	<i>n=1791</i>		

*
3 degrees of freedom

Table 3b

Trail Making Test-B qualitative data [mean (standard deviation)] stratified by educational attainment

Education (degree)	<High School	High School	College	Graduate	Total	F-value*	P-value
Completion time	147.8(68.1)	98.8(52.2)	82.7(41.8)	74.4 (31.3)	91.8(49.1)	49.7	<0.0001
Total errors	1.5(1.7)	0.7(1.1)	0.5(0.9)	0.4(0.7)	0.6(1.0)	18.9	<0.0001
Log Pen lifts	2.1(2.3)	1.6(2.4)	1.2(2.1)	1.0(1.5)	1.4(2.2)	8.6	<0.0001
	<i>n=47</i>	<i>n=1005</i>	<i>n=395</i>	<i>n=344</i>	<i>n=1791</i>		

* 3 degrees of freedom

Table 4a

Verbal Fluency qualitative data [mean (standard deviation)] stratified by age

Age (years)	<55	55-64	65-74	75	Total	F-value*	P-value
FAS							
Total responses	41.6(12.6)	41.8(12.3)	36.3(11.9)	33.4(11.2)	38.0(12.4)	51.8	<0.0001
%Psv errors	2.2(3.5)	3.6(4.2)	3.8(4.7)	5.0(6.3)	3.9(4.9)	13.7	<0.0001
%Total errors	4.3(5.7)	5.7(6.2)	6.3(6.9)	8.3(9.8)	6.4(7.5)	15.6	<0.0001
Animals							
Total responses	20.2(4.1)	20.2(4.9)	17.7(4.7)	14.8(4.5)	18.1(5.1)	125.3	<0.0001
%Psv errors	2.7(4.8)	2.7(7.5)	3.3(5.7)	4.1(8.4)	3.2(7.0)	3.7	0.01
%Total errors	2.7(4.8)	2.9(7.8)	3.6(6.1)	4.6(9.0)	3.5(7.4)	5.2	0.001
	<i>n=132</i>	<i>n=659</i>	<i>n=635</i>	<i>n=429</i>	<i>n=1855</i>		

Note: %Psv errors=perseverations/total responses. %Total errors=total errors/total responses.

*
3 degrees of freedom

Table 4b

Verbal Fluency qualitative data [mean (standard deviation)] stratified by educational attainment

Education (degree)	<High School	High School	College	Graduate	Total	F-value*	P-value
FAS							
Total responses	27.6(13.3)	35.4(11.5)	40.7(12.0)	44.2(11.8)	38.0(12.4)	73.2	<0.0001
%Psv errors	3.8(4.9)	4.2(5.4)	3.3(4.1)	3.6(4.4)	3.9(4.9)	4.1	0.006
%Total errors	9.6(9.9)	7.1(8.2)	5.2(5.8)	5.4(5.9)	6.4(7.5)	12.3	<0.0001
Animals							
Total responses	14.3(5.2)	17.1(4.7)	19.0(5.0)	20.6(5.2)	18.1(5.1)	60.6	<0.0001
%Psv errors	3.6(5.3)	3.7(7.7)	2.4(5.0)	2.7(6.8)	3.2(7.0)	4.1	0.007
%Total errors	3.6(5.3)	4.0(8.1)	2.6(5.6)	3.0(7.2)	3.5(7.4)	4.1	0.007
	<i>n=58</i>	<i>n=1046</i>	<i>n=398</i>	<i>n=353</i>	<i>n=1855</i>		

Note: % Psv errors=perseverations/total responses. %Total errors=total errors/total responses.

*
3 degrees of freedom

Table 5a

WAIS Similarities qualitative data [mean (standard deviation)] stratified by age

Age (years)	<55	55-64	65-74	75	Total	F-value*	P-value
Total score	18.8(2.8)	17.9(3.4)	16.6(3.5)	14.9(4.1)	16.8(3.8)	74.8	<0.0001
Set loss errors	0.4(0.8)	0.6(0.9)	0.7(1.1)	0.9(1.3)	0.7(1.1)	14.2	<0.0001
Concrete responses	4.1(1.5)	4.2(1.7)	4.4(1.8)	4.7(1.8)	4.4(1.8)	8.8	<0.0001
	<i>n=134</i>	<i>n=666</i>	<i>n=646</i>	<i>n=452</i>	<i>n=1898</i>		

*
3 degrees of freedom

Table 5b

WAIS Similarities qualitative data [mean (standard deviation)] stratified by educational attainment

Education (degree)	<High School	High School	College	Graduate	Total	F-value*	P-value
Total score	11.6(4.1)	15.9(3.6)	18.0(3.0)	19.2(2.8)	16.8(3.8)	151.3	<0.0001
Set loss errors	1.3(1.6)	0.8(1.1)	0.6(1.0)	0.5(0.8)	0.7(1.1)	14.1	<0.0001
Concrete responses	5.3(2.3)	4.5(1.8)	4.2(1.7)	3.9(1.5)	4.4(1.8)	20.3	<0.0001
	<i>n=64</i>	<i>n=1073</i>	<i>n=402</i>	<i>n=359</i>	<i>n=1898</i>		

* 3 degrees of freedom

Table 6

Association between brain MRI measures and neuropsychological test measures.

Outcome Variable	Total cerebral brain volume		Frontal lobe volume		Log WMH-L volume		Silent Cerebral Infarct (1)	
	<i>t</i> -value	Beta(SE)	<i>t</i> -value	Beta(SE)	<i>t</i> -value	Beta(SE)	<i>t</i> -value	Beta(SE)
<i>Trailmaking Test-B</i>								
Completion time	-3.2	-1.18(0.37)**	-4.9	-1.95(0.40)***	1.9	2.13(1.12)	0.6	1.81(3.21)
Total errors	-0.4	-0.0036(0.0087)	-2.9	-0.027(0.0094)**	2.7	0.071(0.026)**	-0.6	-0.044(0.076)
Log Pen lifts	-1.3	-0.030(0.023)	-2.7	-0.066(0.025)**	-0.4	0.0030(0.0069)	1.0	0.20(0.20)
<i>FAS</i>								
Total response	2.0	0.20(0.098)*	2.3	0.24(0.11)*	-0.5	-0.14(0.30)	-0.2	-0.21(0.86)
Total errors	-0.6	-0.013(0.020)	0.6	0.012(0.022)	0.8	0.046(0.061)	1.4	0.24(0.17)
Psv errors	0.4	0.0063(0.015)	1.1	0.018(0.017)	0.8	0.037(0.047)	1.8	0.24(0.13)
<i>Animal Naming</i>								
Total response	1.5	0.059(0.040)	1.3	0.057(0.044)	-1.2	-0.14(0.12)	0.2	0.079(0.35)
Total errors	-0.2	-0.0013(0.0078)	0.3	0.0025(0.0085)	0.7	0.018(0.024)	-0.04	-0.0029(0.068)
Psv errors	-0.1	-0.0010(0.0075)	0.2	0.0014(0.0082)	0.7	0.016(0.023)	-0.4	-0.027(0.067)
<i>Similarities</i>								
Total score	3.1	0.087(0.028)**	1.7	0.052(0.031)	0.4	0.035(0.086)	1.2	0.30(0.25)
Concrete errors	-0.9	-0.014(0.015)	-1.0	-0.016(0.017)	0.5	0.021(0.046)	-0.7	-0.091(0.13)
	<i>Chi-square</i>	<i>RR(95% CI)</i>	<i>Chi-square</i>	<i>RR(95% CI)</i>	<i>Chi-square</i>	<i>RR(95% CI)</i>	<i>Chi-square</i>	<i>RR(95% CI)</i>
Set maintenance errors (1)	1.5	0.98(0.95-1.01)	0.4	1.01(0.97-1.05)	0.6	1.04(0.94-1.16)	1.4	0.83(0.60-1.13)

Note: All models are adjusted for age, sex, and education group. All statistical tests use 1 degree of freedom. WMH-L, Large white matter hyperintensities. TrB, Trailmaking Test-part B. Psv, perseveration.

* p<0.05,

** p<0.01,

*** p<0.001

Table 7a

Neuropsychological battery

Memory, Attention, Visuospatial & Language Tests
WMS Logical Memory: Immediate, Delayed, Recognition
WMS Visual Reproduction: Immediate, Delayed, Recognition
WMS Paired Associate Learning: Immediate, Delayed, Recognition
WAIS Digit Span
Trailmaking Test - part A
Boston Naming Test, 36 items
Hooper Visual Organization Test

Note: WMS=Wechsler Memory Scale. WAIS=Wechsler Adult Intelligence Scale.

Table 7b

Qualitative errors on neuropsychological test battery

Domain	Test	Error type	Range of errors
Memory	LM Immediate, Delayed	Intrusion errors	0 – max
		Non-response	0 – 11
	VR Immediate, Delayed	Confabulation errors	0 – 4
		Perseveration errors	0 – 4
	VR Recognition	Non-response	0 – 4
	VPA Immediate, Delayed	Interference errors	0 – 10 (per trial)
		Intrusion errors	0 – 10 (per trial)
		Perseveration errors	0 – 10 (per trial)
VPA Recognition	Non-response	0 – 10	
Attention	Digit span	Sequencing errors	0 – 14
		Non-sequencing errors	0 – 8
	Trail-making test-A	Qualitative errors	0 – max
Language	Boston Naming Test	Circumlocution	0 – 36
		Perseveration errors	0 – 35
		Paraphasic errors	0 – 36
		Perceptual errors	0 – 36
Visuospatial	HVOT	Isolate errors	0 – 30
		Perceptual errors	0 – 30

Note: LM=WMS Logical Memory. VR=WMS Visual Reproduction. VPA=WMS Verbal Paired Associates. HVOT=Hooper Visual Organization Test.

Table 7c

WMS Logical Memory qualitative data [mean (standard deviation)] stratified by age and education

Age, years (n)	<55 (135)	55-64 (666)	65-74 (647)	75 (456)
Immediate–Intrusions	1.1 (1.2)	1.1 (1.1)	1.2 (1.3)	1.2 (1.2)
Delayed–Intrusions	1.7 (1.6)	1.6 (1.4)	1.8 (1.5)	1.7 (1.5)
Recognition–Non-response	0.0(0.2)	0.1(0.3)	0.1(0.4)	0.1(0.5)
Education, degree (n)	<High School (64)	High School (1076)	College (404)	Graduate (360)
Immediate–Intrusions	0.9 (1.1)	1.2 (1.2)	1.2 (1.1)	1.2 (1.2)
Delayed–Intrusions	1.0 (1.3)	1.7 (1.5)	1.9 (1.5)	1.8 (1.5)
Recognition–Non-response	0.2(0.6)	0.1(0.4)	0.0(0.2)	0.1(0.3)

Table 7d

WMS Visual Reproductions qualitative data [mean (standard deviation)] stratified by age and education

Age, years (n)	<55 (135)	55-64 (666)	65-74 (645)	75 (446)
Immediate-Confabulation	0.1 (0.4)	0.1 (0.4)	0.2 (0.5)	0.4 (0.7)
Immediate-Perseveration	0.1(0.3)	0.2(0.4)	0.3(0.5)	0.4(0.6)
Delayed-Confabulation	0.1 (0.3)	0.2 (0.5)	0.2 (0.6)	0.3 (0.6)
Delayed-Perseveration	0.1(0.3)	0.2(0.4)	0.3 (0.5)	0.3(0.6)
Recognition-Non-response	0.0(0.0)	0.0(0.1)	0.0(0.1)	0.0(0.2)
Education, degree (n)	<High School (63)	High School (1068)	College (403)	Graduate (358)
Immediate-Confabulation	0.4 (0.6)	0.3 (0.6)	0.2 (0.5)	0.2 (0.5)
Immediate-Perseveration	0.3(0.5)	0.3(0.5)	0.2(0.4)	0.2(0.5)
Delayed-Confabulation	0.3 (0.7)	0.2 (0.6)	0.2 (0.5)	0.1 (0.4)
Delayed-Perseveration	0.3(0.5)	0.3(0.5)	0.2(0.4)	0.2(0.5)
Recognition-Non-response	0.0(0.2)	0.0(0.2)	0.0(0.1)	0.0(0.1)

Table 7e

WMS Verbal Paired Associates qualitative data [mean (standard deviation)] stratified by age and education

Age, years (n)	<55 (134)	55-64 (664)	65-74 (643)	75 (447)
Immediate–Interference errors	2.3(2.1)	2.2(2.0)	2.3(2.0)	2.0(1.9)
Immediate–Intrusion errors	0.7 (1.1)	0.9 (1.3)	1.0 (1.4)	1.6 (2.1)
Immediate–Perseveration errors	0.1(0.4)	0.2(0.5)	0.3(0.6)	0.3(0.8)
Delayed–Interference errors	0.7(0.9)	0.8(0.9)	0.8(0.9)	0.7(0.9)
Delayed–Intrusion errors	0.7 (1.0)	0.8 (1.0)	0.9 (0.1)	1.0 (1.1)
Delayed–Perseveration errors	0.2(0.5)	0.2(0.5)	0.3(0.6)	0.4(0.7)
Recognition–Non-response	0.0(0.1)	0.0(0.1)	0.0(0.2)	0.1(0.3)
Education, degree (n)	<High School (61)	High School (1065)	College (404)	Graduate (358)
Immediate–Interference errors	1.8(1.6)	2.3(2.0)	2.0(1.9)	2.1(2.0)
Immediate–Intrusions	2.1 (2.3)	1.1 (1.6)	1.0 (1.4)	0.8 (1.2)
Immediate–Perseveration errors	0.3(1.1)	0.3(0.6)	0.2(0.5)	0.2(0.6)
Delayed–Interference errors	0.8(0.9)	0.9(1.0)	0.7(0.9)	0.6(0.9)
Delayed–Intrusions	1.0 (1.1)	1.0 (1.1)	0.8 (1.0)	0.6 (0.9)
Delayed–Perseveration errors	0.4(0.8)	0.3(0.6)	0.2(0.5)	0.2(0.5)
Recognition–Non-response	0.1(0.2)	0.0(0.2)	0.0(0.1)	0.0(0.1)

Table 7f

WAIS Digit Span qualitative data [mean (standard deviation)] stratified by age and education

Age, years (n)	<55 (123)	55-64 (619)	65-74 (604)	75 (426)
Forward-Tests administered	8.5(1.2)	8.5(1.4)	8.4(1.4)	8.1(1.5)
Forward-Errors	3.1(1.4)	3.4(1.5)	3.5(1.5)	3.5(1.4)
Forward-Sequencing errors	0.9(0.9)	1.0(1.1)	1.0(1.0)	1.0(1.1)
Forward-% Errors	36.5(14.3)	39.4(14.8)	40.5(14.3)	42.6(13.9)
Forward-% Sequencing errors	26.2(24.3)	25.6(23.9)	24.8(22.6)	23.3(22.4)
Backward-Tests administered	6.6(1.6)	7.2(1.7)	7.5(1.7)	7.6(1.6)
Backward-Errors	3.6(1.5)	3.4(1.4)	3.3(1.4)	3.2(1.4)
Backward-Sequencing errors	1.1(1.2)	0.9(1.0)	0.8(1.0)	0.8(0.9)
Backward-% Errors	53.6(16.7)	46.3(13.1)	43.5(13.4)	41.8(13.8)
Backward-% Sequencing errors	22.6(22.8)	20.3(21.6)	18.0(20.6)	20.0(20.2)
Education, degree (n)	<High School (58)	High School (1004)	College (380)	Graduate (330)
Forward-Tests administered	8.0(1.7)	8.3(1.5)	8.4(1.3)	8.5(1.3)
Forward-Errors	3.6(1.5)	3.5(1.5)	3.3(1.5)	3.2(1.5)
Forward-Sequencing errors	1.0(1.1)	1.1(1.1)	1.0(1.0)	0.9(1.0)
Forward-% Errors	44.6(14.3)	41.9(14.2)	38.6(14.1)	36.7(14.9)
Forward-% Sequencing errors	21.9(21.4)	25.6(23.3)	24.0(22.3)	23.8(24.1)
Backward-Tests administered	6.6(1.6)	7.2(1.7)	7.5(1.7)	7.6(1.6)
Backward-Errors	3.6(1.5)	3.4(1.4)	3.3(1.4)	3.2(1.4)
Backward-Sequencing errors	1.1(1.2)	0.9(1.0)	0.8(1.0)	0.8(0.9)
Backward-% Errors	53.6(16.7)	46.3(13.1)	43.5(13.4)	41.8(13.8)
Backward-% Sequencing errors	22.6(22.8)	20.3(21.6)	18.0(20.6)	20.0(20.2)

Note: % Error responses=Errors/Tests administered. % Sequencing errors=Sequencing errors/Tests administered.

Table 7g

Trail Making Test – Part A qualitative data [mean (standard deviation)] stratified by age and education

Age, years (n)	<55 (134)	55-64 (659)	65-74 (639)	75 (422)
Errors	0.1(0.3)	0.1(0.4)	0.1(0.4)	0.2(0.4)
Pen lifts	0.5(0.7)	0.7(1.0)	1.0(1.6)	1.6(2.2)
Early start	0.1(0.3)	0.1(0.3)	0.2(0.4)	0.2(0.4)
Education, degree (n)	<High School (62)	High School (1043)	College (398)	Graduate (351)
Errors	0.1(0.4)	0.2(0.4)	0.1(0.4)	0.1(0.4)
Pen lifts	1.4(1.5)	1.1(1.6)	0.9(1.9)	0.8(1.2)
Early start	0.3(0.5)	0.2(0.4)	0.1(0.4)	0.1(0.4)

Table 7h

Boston Naming Test qualitative data [mean (standard deviation)] stratified by age and education

Age, years (n)	<55 (135)	55-64 (666)	65-74 (645)	75 (441)
Circumlocution errors	0.5(1.1)	0.7(1.3)	1.1(2.0)	1.5(2.1)
Perseveration errors	0.0(0.1)	0.0(0.2)	0.1(0.3)	0.1(0.4)
Semantic paraphasic errors	1.1(1.3)	1.4(1.6)	1.7(1.7)	2.1(1.9)
Phonemic paraphasic errors	0.2(0.5)	0.3(0.7)	0.5(0.9)	0.7(1.0)
Perceptual errors	0.2(0.4)	0.2(0.5)	0.3(0.7)	0.5(1.0)
Education, degree (n)	<High School (64)	High School (1066)	College (401)	Graduate (356)
Circumlocution errors	1.7(2.0)	1.2(2.0)	0.9(1.6)	0.5(1.1)
Perseveration errors	0.1(0.4)	0.1(0.3)	0.0(0.3)	0.1(0.3)
Semantic paraphasic errors	2.1(1.6)	1.8(1.9)	1.4(1.5)	1.3(1.6)
Phonemic paraphasic errors	1.0(1.2)	0.5(1.0)	0.3(0.6)	0.2(0.6)
Perceptual errors	0.5(0.8)	0.4(0.8)	0.3(0.7)	0.2(0.7)

Table 7i

Hooper Visual Organization Test qualitative data [mean (standard deviation)] stratified by age and education

Age, years (n)	<55 (132)	55-64 (650)	65-74 (624)	75 (424)
Isolate errors	0.8(1.2)	1.0(1.3)	1.2(1.5)	1.5(1.6)
Perceptual errors	1.2(1.3)	1.3(1.6)	1.5(1.8)	1.8(1.9)
Education, degree (n)	<High School (59)	High School (1050)	College (400)	Graduate (351)
Isolate errors	1.5(1.6)	1.2(1.5)	1.1(1.4)	1.0(1.4)
Perceptual errors	1.7(1.6)	1.6(1.8)	1.5(1.6)	1.3(1.6)