

## **Neuropsychological Test Performance of Cognitively Healthy Centenarians: Normative data from the Dutch 100-plus Study**

**Authors:** Nina Beker, MSc<sup>1,2</sup>, Sietske A.M. Sikkes, PhD<sup>1,2,4</sup>, Marc Hulsman, PhD<sup>2,3</sup>, Ben Schmand, PhD<sup>5</sup>, Philip Scheltens, MD, PhD<sup>1,2</sup> and Henne Holstege, PhD<sup>1,2,3</sup>

### **Affiliations:**

<sup>1</sup> VU University Medical Center, Alzheimer Center & Department of Neurology, Amsterdam, the Netherlands

<sup>2</sup> Amsterdam Neuroscience, Amsterdam, the Netherlands

<sup>3</sup> VU University Medical Center, Department of Clinical Genetics, Amsterdam, the Netherlands

<sup>4</sup> Massachusetts General Hospital, Department of Neurology, Boston, USA

<sup>5</sup> University of Amsterdam, Department of Psychology, Amsterdam, the Netherlands

**Corresponding author:** Henne Holstege, Alzheimer Center & Department of Neurology, VU University Medical Center, PO Box 7057, 1007 MB Amsterdam, the Netherlands (h.holstege@vumc.nl).

**Keywords:** Normative data, cognitive functioning, extreme ageing, oldest-old, neuropsychological tests, cognitive health, 100-plus Study

## ABSTRACT

**Background:** The population who reaches the extreme age of 100 years is growing. At this age, dementia incidence is high and cognitive functioning is variable and influenced by sensory impairments. Appropriate cognitive testing requires normative data generated specifically for this group. Currently, these are lacking. We set out to generate norms for neuropsychological tests in cognitively healthy centenarians while taking sensory impairments into account.

**Methods:** We included 235 centenarians (71.5% female) from the 100-plus Study, who self-reported to be cognitively healthy, which was confirmed by an informant and a trained researcher. Normative data were generated for 15 tests that evaluate global cognition, pre-morbid intelligence, attention, language, memory, executive and visuo-spatial functions by multiple linear regressions and/or percentiles. Centenarians with vision and/or hearing impairments were excluded for tests that required these faculties.

**Results:** Subjects scored on average  $25.6 \pm 3.1$  (range 17-30, interquartile-range 24-28) points on the MMSE. Vision problems and fatigue often complicated the ability to complete tests, and these problems explained 41% and 22% of the missing test scores respectively, whereas hearing problems (4%) and task incomprehension (6%) only rarely did. Sex and age showed a limited association with test performance, whereas educational level was associated with performance on the majority of the tests.

**Conclusions:** Normative data for the centenarian population is provided, while taking age-related sensory impairments into consideration. Results indicate that, next to vision impairments, fatigue and education level should be taken into account when assessing cognitive functioning in centenarians.

## INTRODUCTION

Worldwide, the number of centenarians is expected to increase almost 20-fold to 3.2 million people in the next 30 years [1]. The incidence of dementia increases exponentially with age, reaching ~40% at age 100 [2]. Previous studies estimated that 25% of centenarians show no objective evidence of memory deficits, while 25% have symptoms of cognitive impairment and 50% may be regarded as having dementia [3, 4]. To evaluate cognitive impairment in this heterogeneous group, it is of great importance that suitable instruments be implemented that take into account the specific characteristics that are observed in centenarians.

Most commonly, evaluation of cognitive functioning in centenarians is based on normative data generated in younger elderly. However, the use of cut-off scores derived in younger samples may not account for the cognitive decline as part of the normal ageing process. Therefore, when these norms are used to assess cognitive performance of a centenarian, this may lead to misclassifications of cognitive impairment [5]. Several studies showed that specific cognitive functions decline during the process of ageing, including a slower processing speed and a decline in executive functions, whereas domains such as language and semantic knowledge are relatively maintained [6-8]. Between elderly aged 80 and 100 years, significant differences in cognitive test performance across cognitive domains were observed with increasing age [5, 9, 10]. This effect culminated in centenarians; their mean performance on cognitive test performance was lowest, while the variability in test performance was increased compared to 80 and 90 year olds [11, 12]. These findings suggest that accurate assessments of cognitive functioning of elderly at extreme ages can only be obtained using normative data generated in cohorts with relatively narrow age bands [9].

Thus far, most studies that provided normative data for centenarians described their performance on the Mini-Mental State Examination (MMSE) [12-16]. The MMSE is used to screen global cognitive functioning, while additional cognitive tests are required to evaluate a broad spectrum of cognitive domains. The Georgia Centenarian Study (GCS) provided data on tests that evaluate multiple cognitive domains, among which the Controlled Oral Word Association Test, the Fuld Object Memory Evaluation and Severe Impairment Battery and the Behavioral Dyscontrol Scale [12, 17-19]. However, these data (i) were generated in population-based samples, which included centenarians with cognitive impairment and (ii) were not adjusted for loss of hearing and sight. This will lead to lower norm-ranges, which might complicate the evaluation of cognitive health in a clinical setting. To evaluate cognitive performance in centenarians, it is necessary to compare their test performance with the performance of a large cognitively intact sample, and to take into account vision and hearing disabilities. The 90+ Study previously included non-demented subjects in the age bands (90-91, 92-94, 95+) to provide suitable normative data [10]. Despite the test adaptations made to compensate for sensory losses in the 90+ Study, some subjects could not complete tests due to sensory impairments, which emphasizes the importance of taking these faculties into account when establishing normative data.

Here, we aim to generate normative data for the evaluation of cognitive functioning in centenarians. We used a large sample of Dutch cognitively healthy centenarians from the 100-plus Study [20]. Per test, we considered whether vision and hearing abilities should be taken into account. The cognitive tests used in this study cover a wide range of cognitive domains including global cognition, pre-morbid intelligence, attention, language, memory, executive and visuo-spatial functioning.

## **METHODS**

### **Subjects**

Subjects were part of the Dutch 100-plus Study cohort, a longitudinal cohort study of people who 1) were aged  $\geq 100$  years, 2) self-reported to be cognitively healthy, which was confirmed by family members and/or caregivers. For this study, we specifically selected subjects who were estimated to be cognitively healthy by a trained researcher and of whom we had at least one neuropsychological test available. We excluded subjects with vision and/or hearing difficulties for tests requiring these sensory functions, leading to a total sample of  $N=235$ . A description of inclusion procedures can be found elsewhere [20]. This study was approved by the Medical Ethics Committee and all subjects provided informed consent.

### **Procedure**

#### *Estimation of cognitive health, hearing and vision*

The subjects were visited at home by researchers with a neuropsychological and/or medical training. The researchers noted their clinical impression of the cognitive health of the centenarians. This estimation was based on features such as centenarians continually repeating themselves, having difficulty understanding or remembering questions, and with naming and/or word finding. During study visits, vision and hearing were subjectively evaluated by the researcher and categorized into “good”, “moderate”, “poor” and “very poor” based on the features listed in the supplemental material. Subjects with poor to very poor vision were excluded from reporting normative data of the Mini-Mental State Examination (MMSE), the Key Search test, the Dutch Adult Reading Test (DART), the Visual Association Test (VAT), Trail Making Test (TMT), Number Location and the Clock Drawing Test (CDT). Subjects with poor to very poor hearing were excluded from reporting normative data on the MMSE, Digit Span and the Rivermead Behavioural Memory Test (RBMT).

### *Neuropsychological testing*

We encouraged the presence of a close relation during study visits. Close relations were requested to not interrupt during the administration of the cognitive tests. The test battery took approximately 1.5 hours to complete, and we took short breaks whenever subjects showed signs of fatigue. Subjects were encouraged to use all available devices to support their vision and/or hearing. Tests were aborted when sensory problems clearly interfered with test performance. When tests could not be administered or needed to be aborted, we annotated the reasons for interference with test completion: physical, vision or hearing problems, fatigue, or incapable of understanding tasks or instructions.

## **Measures**

### *Neuropsychological testing*

The neuropsychological test battery consisted of fifteen well-known tests that evaluate global cognition, pre-morbid intelligence, attention, language, memory, executive and visuo-spatial functions. The test battery has expanded gradually over the course of the study according to new insights and experience with testing centenarians. Originally, the battery was limited to short screening tests including the MMSE and the CDT and it expanded over time with tests that evaluate specific cognitive domains. Accordingly, not all centenarians were presented with the same test battery or the same number of tests. In total, 77% (n=180) of the centenarians were subjected to a test battery including all 15 tests evaluated here. Tests were administered according to standard procedures, unless stated otherwise.

### *Mini-Mental State Examination*

The Dutch version of the MMSE [21, 22] was administered to evaluate global cognitive functioning. The MMSE covers multiple cognitive domains including orientation, memory, attention and/or concentration, working memory, language, visual construction and praxis. The score range for the MMSE is 0 – 30, with higher scores indicating better performance.

### *National Adult Reading Test*

We addressed pre-morbid intelligence using the DART, a Dutch version of the National Adult Reading Test [23-25]. Subjects were asked to read out loud fifty words with atypical phonemic pronunciation. We presented these words in a large font-size because of the frequent visual limitations in centenarians.

### *WAIS Digit Span Forward and Backward*

We used the Digit Span Forward of the Wechsler Adult Intelligence Scale-III to evaluate attention and/or concentration and the Digit Span Backward to measure working memory [26]. The Forward condition involves subjects to repeat sequences of digits, that increase in length. In the Backward condition, subjects are asked to repeat the increasing digit sequences in reverse order. For both the Digit Span Forward and Backward tests, we present norms for the total raw scores (range 0-16) and the total span of digits (range 0-8), with higher scores indicating better performance.

### *Letter and Category Fluency*

The Dutch version of the Controlled Oral Word Association Test (COWAT) [27, 28] was administered to evaluate executive functioning and language. In this Letter Fluency task, subjects were asked to name as many words starting with the letter “D”, “A”, and “T”, each in 1 minute. We also used the Animal Fluency (naming animals in 1 and 2 minutes) to evaluate executive functioning, language and semantic memory [29, 30].

### *Rivermead Behavioural Memory Test*

Episodic memory was measured with the story recall subtest of the Dutch version of the RBMT [31-33]. The RBMT contains news articles, which are read out loud to the subjects. The subjects were asked to name all story-items they remembered (immediate recall) and this was repeated after a 15-minute interval (delayed recall). When necessary, the subjects were given a cue for helping them recall the story-line; this cue was not included when

calculating the total score. We made two adaptations to the test procedure: 1) two articles were read instead of one in order to improve reliability; 2) during recall all remembered items were scored, whether they belonged to the appropriate story-line or not. The score range is 0-42 for both immediate and delayed recall sections, with higher scores indicating better performance.

#### *Visual Association Test*

The VAT was developed to detect anterograde amnesia and is based on paired associated learning [34, 35]. Subjects are asked to name two interacting visual items (object or animals, e.g. a hedgehog on a chair), of which one item (hedgehog) needs to be recalled afterwards while the other (chair) is used as a cue. We used naming of the items as an additional measure of language functioning. The score range is 0-12, with higher scores indicating better performance.

#### *Trail Making Test*

We used TMT A for evaluating processing speed and attention, and TMT B to measure mental flexibility [36]. TMT A requires the subjects to connect 15 dots of numbers in numerical order. In TMT B, subjects were asked to alternate between numbers and letters in numerical and alphabetical sequence (1-A-2-B etc.). For TMT A after 180 seconds, and for TMT B after 300 seconds, we ascertained whether the subjects were still determined to proceed. If this not was the case the test was aborted and scores were extrapolated based on the last finished item (number or letter) and the time spent on the test.

#### *BADS Key Search*

The Behavioral Assessment of the Dysexecutive Syndrome Test Battery (BADS) [37] was developed to measure executive functioning. Here we used the Key Search subtest, which involves a problem-solving task instructing subjects to think of a strategy to find a lost key.



Scores range from 0-16, with higher scores indicating better performance. A time limit was not imposed.

#### *VOSP Number location*

Visuo-spatial orientation was assessed with the Number Location subtest of the Visual Object and Space Perception Battery (VOSP) [38], in which subjects were asked to visualize and point out a specific number that corresponded with the exact location of a dot. Scores range from 0-10, with higher scores indicating better performance.

#### *Clock drawing test*

Visuo-spatial construction was measured by the Clock Drawing Test (CDT) [39, 40], in which subjects were instructed to draw a clock with the hands at 10 past 11. Because of poor pen-holding skills or tremor in our subjects, we offered a pre-drawn circle [41]. The scores range from 0 to 5, whereas a score of 0 was given in case of no reasonable representation of a clock, and 5 indicated a perfect clock [40]. In this scoring method, the drawing of the circle itself was not considered.

#### **Demographic and clinical measures**

Education level was determined based on the International Standard Classification of Education 1997 (ISCED) [42]. Independence in performing activities of daily living (ADL) was evaluated with the Dutch observation and/or interview version of the Barthel Index [43-46]. Scores ranged from 0-20, where scores of <15 indicate dependence, and scores  $\geq 15$  indicate independence in performing ADL [46]. The 15-item version of the Geriatric Depression scale (GDS) was administered to investigate depressive symptoms [47]. Scores ranged from 0-15:  $\leq 5$  indicates no depressive symptoms and  $> 5$  suggests the presence of depressive symptoms.

## **Data analysis**

Normative data were generated by multiple linear regression analyses and by reporting means and percentiles.

The multiple linear regression analyses with age, sex, and education as independent variables and cognitive test scores as dependent variables were performed for the cognitive tests separately. For the regression analyses, TMT scores were log-transformed because of non-normal distributions and inverted, so that higher scores indicated better performance. Because of ceiling effects, the VAT, Number Location and the CDT were not included in the regression analyses. The means and percentiles were reported for each test for the total sample and stratified by education level: low education (post-secondary non-tertiary or lower) and high education (short-cycle tertiary or higher). The cognitive test scores were standardized into z-scores in order to I) correlate the number of tests the centenarians were able to complete with their overall mean z-score and II) to visualize the distribution of the test scores using boxplots. P-values <0.05 were considered significant. Statistical analyses were performed using SPSS 22.0 and R 3.4.2.

## RESULTS

### Demographic and clinical characteristics

The total sample of N=235 had a median age of 100.4 years (range 100-107) at inclusion and included 168 females (71%). The majority (59%) of the centenarians lived independently, 79% were independently mobile and 54% were independent in performing ADL. Most of the centenarians had retained moderate-good vision (77%) and hearing capacities (88%). The majority of the centenarians (62%) had a basic-low education level. Most centenarians (92%) did not show depressive symptoms as measured with the GDS. Clinical and demographic characteristics for this group are summarized in Table 1.

### Influence of age-related sensory impairments on test incompleteness

Across all tests, an average of 79% of the tests was completed (Table 2). While >95% of the subjects completed both fluency tasks, only 45% were able to complete TMT B. Difficulties with vision (41%) and fatigue (22%) were the most common reasons for not being able to complete a test, whereas hearing impairment only rarely complicated test completion (4%). In some cases, not understanding the test and/or test instructions was a reason for incompleteness of the Number Location (16%), Key Search (14%), and TMT B (23%). Overall, we found a positive correlation between the number of tests a centenarian was able to complete and the mean z-score across all completed tests (Pearson's correlation,  $r=.35$ ,  $p<.001$ ), see supplemental data.

### Normative data and cognitive test performance across centenarians

Per test, an overview of the sample-sizes used for generating normative data is shown in Table 2. Centenarians with poor to very poor vision (21%) and hearing (11%) were excluded for tests for which these faculties were required. We present percentiles and means of all cognitive tests stratified by educational level (Table 3). In the total sample, the mean  $\pm$ SD score on the MMSE was  $25.6 \pm 3.1$  (range 17-30, interquartile-range 24-28). High-educated centenarians scored on average  $26.5 \pm 3.0$ , while the centenarians who attained lower

education levels scored lower ( $25.2 \pm 3.1$ ). In addition, multiple linear regression analyses were performed to provide normative data adjusted for sex, age, and education (Table 4). For clinical utility, we also provide the regression formulas for all tests (supplementary material). Correlations between the cognitive test scores are displayed in the supplement. Figure 1 shows the distribution of the performances on each test. Overall, most test scores showed wide distributions, while the VAT, Number Location and the CDT had strong ceiling effects (Table 3).

### **Association of education, sex, and age with cognitive test performance**

The level of educational attainment was positively associated with the performance on all tests when adjusted for age and sex, except the RBMT Delayed Recall, TMT A and B. Sex differences were found for the Digit Span Forward, Key Search and Letter Fluency, with males having higher test scores on the Digit Span Forward ( $mean \pm SD = 7.7 \pm 1.8$  versus females  $6.9 \pm 1.8$ ), and the Key Search ( $mean = 8.1 \pm 4.1$  vs  $mean = 6.1 \pm 3.3$ ). Females had higher test scores on the Letter Fluency ( $mean = 25.0 \pm 10.4$ ) than males ( $mean = 22.9 \pm 10.6$ ). On all other tests, males and females performed similarly. Age was only associated with the performance on the TMT A when adjusted for sex and education. The results of these analyses are presented in Table 4.

## **DISCUSSION**

A large sample of cognitively healthy centenarians was used to generate normative data of fifteen common neuropsychological tests across several cognitive domains, whilst taking vision and hearing impairments into account. We found that test completion was most often complicated by vision impairments and fatigue, and to a lesser extent by hearing impairments or task incomprehension. Almost all cognitive test scores were associated with educational attainment.

### **Cognitive test performance in centenarians**

Previous studies reported that the variability in cognitive test performance increased with age [11, 12]. Accordingly, most test scores showed wide distributions in the current study, indicating that cognitive test performance was variable within the centenarians. For the VAT, Number Location and the CDT we observed ceiling effects, suggesting that these tests are relatively easy to complete, and might be limited in their ability to capture cognitive performance for this specific population. In line with previous studies, we found that some subjects had difficulty to complete executive functioning tests, supporting the theory that executive functioning is particularly vulnerable to decline in normal ageing [6, 10]. In contrast, the fluency tests were completed by almost all centenarians with varying results and without floor- and ceiling effects, implying that these are suitable tests for application to the oldest-old.

As expected, this selected sample of centenarians scored on average  $25.6 \pm 3.1$  points on the MMSE, which is above the suggested cut-off score used for indicating cognitive impairment in elderly aged 97 years and above [48]. This score is similar to the mean MMSE score of  $\sim 24$  obtained by American cognitively intact centenarians [15], and markedly higher than the scores of centenarians from population-based studies, which varied from 12.5 to 20 [12-14, 49-51]. More specifically, the test performance of the centenarians on the MMSE, Animal and Letter Fluency, Digit Span Backward, and TMT B, was comparable to scores of the non-

demented subjects aged >95 years from the 90+ Study, while performance of the centenarians on Digit Span Forward and TMT A was lower [10]. This could indicate that processing speed and attention may decline in the years between 95 and 100+, while other cognitive domains remain stable. Further, despite selection against vision and hearing impairments in the sample of centenarians, the difference in cognitive test performance might be explained by their increased vulnerability for sensory decline.

### **Influence of age-related sensory impairments on test incompleteness**

Overall, we found an association between the ability to complete tests and test performance, which emphasizes the importance of taking into account factors that interfere with test completion when assessing cognitive functions. Our results suggest that visual impairments, but not hearing impairments, were the most common reason for the inability to complete tests, which is in agreement with previous reports [10]. Therefore, we caution that tests that require vision ability may not be fully applicable in centenarians and that there is room for the design of tests that can be applied independently of these vision impairments. Next to visual impairments, also fatigue commonly led to test incompleteness, suggesting that our test battery may have been too extensive for a subset of the centenarians. Therefore, to prevent fatigue from interfering with test results, tests and test batteries for the oldest-old should be kept as short as possible [10, 52].

### **Influence of education, sex and age on cognitive test performance**

Consistent with previous findings, we found education to be associated with performance on almost all tests, except for delayed recall and the TMT [5, 10]. Accordingly, we expect that having a lower educational attainment is representative of the lower range of the normative data. More specifically, previous studies showed that lower educated elderly scored below cut-off scores on cognitive screening tests, which could cause an overestimation of cognitive impairment in these elderly [53-56]. This might explain that some centenarians, while appearing cognitively healthy during study visits, scored <23 points on the MMSE, or had

poor performances on tests on which the majority of the centenarians obtained maximum scores. Therefore, we emphasize that low performance on individual tests should be interpreted in the context of other test scores, taking demographic characteristics and factors of sensory decline into account, to evaluate cognitive impairment in centenarians.

With regard to sex, *population-based* centenarian studies indicated that males consistently performed better on cognitive tests [17, 57], which might reflect the higher dementia prevalence in centenarian females [58, 59]. However, our inclusion criteria of “cognitively healthy centenarians” may introduce a selection bias for only the cognitively healthy males and females, which might explain why we observe no clear sex difference in cognitive test performance. Finally, whereas age is seen as one of the most important predictors for cognitive decline, this was not a strong predictor for cognitive performance in centenarians. We expect that the age range of 100-107, with an interquartile range of 100-101, was too narrow to investigate the effect of age on cognitive test performance.

### **Strengths and limitations**

In this study we generated normative data specifically tailored for assessing cognitive performance in centenarians, which, given their high risk of having or developing symptoms of cognitive impairment is expected to be variable. The availability of a relatively large sample of cognitively healthy centenarians allowed us to exclude centenarians with sensory losses for the generation of normative data for tests that require vision and hearing ability. Since our test battery expanded during the course of the study, not every test was administered to all subjects. Accordingly, the sample sizes for reporting normative data were unequal and consequently, the order of the tests in the battery could not be taken into account when assessing the influence of fatigue. Further, to select a cognitively healthy population we relied on self-reported cognitive health, which was confirmed by us and by an informant, however, subjects were not screened for dementia according to standard NIA-AA criteria. Nevertheless, we find that the centenarians included in the current study are

heterogeneous based on several aspects but based on the MMSE scores, they can be regarded among the highest performing centenarians. Therefore, our cohort of centenarians is appropriate for the generation of normative data to evaluate cognitive function in centenarians.

## **Conclusion**

This study provides normative data on a broad range of cognitive tests in a large sample of cognitively healthy centenarians, taking hearing and vision impairments into account. The norms established in the current study are useful tools for the detection of dementia and cognitive impairment in centenarians. Our results suggest that, besides vision impairment, also fatigue and education level should be considered when evaluating test results in centenarians.



## **Acknowledgements**

We are grateful for the collaborative efforts of all participating centenarians and their family members and/or relations. We acknowledge the persons who visited and/or recruited the centenarians. We thank Sven van der Lee, Niccolò Tesi and Wiesje van der Flier who advised on the analyses and presentation of the normative data. This work was funded by Stichting Alzheimer Nederland (WE09.2014-03), Stichting Dioraphte (VSM 14 04 1402) and Stichting VUmc Fonds. The authors declare no conflict of interest.

## References

1. United Nations, D.o.E.a.S.A. *World Population Ageing: 1950-2050* 2001. 23-27 DOI: ST/ESA/SER.A/207.
2. Corrada, M.M., et al., *Dementia incidence continues to increase with age in the oldest old: the 90+ study*. *Ann Neurol*, 2010. **67**(1): p. 114-21.
3. Peltz, C.B., et al., *Cognitive impairment in nondemented oldest-old: prevalence and relationship to cardiovascular risk factors*. *Alzheimers Dement*, 2012. **8**(2): p. 87-94.
4. Poon, L.W., et al., *Understanding dementia prevalence among centenarians*. *J Gerontol A Biol Sci Med Sci*, 2012. **67**(4): p. 358-65.
5. Beerli, M.S., et al., *Age, gender, and education norms on the CERAD neuropsychological battery in the oldest old*. *Neurology*, 2006. **67**(6): p. 1006-10.
6. Harada, C.N., M.C. Natelson Love, and K.L. Triebel, *Normal cognitive aging*. *Clin Geriatr Med*, 2013. **29**(4): p. 737-52.
7. Hedden, T. and J.D. Gabrieli, *Insights into the ageing mind: a view from cognitive neuroscience*. *Nat Rev Neurosci*, 2004. **5**(2): p. 87-96.
8. Verhaeghen, P. and T.A. Salthouse, *Meta-analyses of age-cognition relations in adulthood: estimates of linear and nonlinear age effects and structural models*. *Psychol Bull*, 1997. **122**(3): p. 231-49.
9. Miller, I.N., et al., *Normative Data for the Cognitively Intact Oldest-Old: The Framingham Heart Study*. *Exp Aging Res*, 2015. **41**(4): p. 386-409.
10. Whittle, C., et al., *Neuropsychological data in nondemented oldest old: the 90+ Study*. *J Clin Exp Neuropsychol*, 2007. **29**(3): p. 290-9.
11. Hagberg, B., et al., *Cognitive functioning in centenarians: a coordinated analysis of results from three countries*. *J Gerontol B Psychol Sci Soc Sci*, 2001. **56**(3): p. P141-51.
12. Miller, L.S., et al., *Cognitive performance in centenarians and the oldest old: norms from the Georgia Centenarian Study*. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn*, 2010. **17**(5): p. 575-90.
13. Engberg, H., et al., *Cohort changes in cognitive function among Danish centenarians. A comparative study of 2 birth cohorts born in 1895 and 1905*. *Dement Geriatr Cogn Disord*, 2008. **26**(2): p. 153-60.
14. Gondo, Y., et al., *Functional status of centenarians in Tokyo, Japan: developing better phenotypes of exceptional longevity*. *J Gerontol A Biol Sci Med Sci*, 2006. **61**(3): p. 305-10.
15. Holtsberg, P.A., et al., *Mini-Mental State Exam status of community-dwelling cognitively intact centenarians*. *Int Psychogeriatr*, 1995. **7**(3): p. 417-27.
16. Inagaki, H., et al., *Cognitive function in Japanese centenarians according to the Mini-Mental State Examination*. *Dement Geriatr Cogn Disord*, 2009. **28**(1): p. 6-12.
17. Davey, A., et al., *Cognitive function, physical performance, health, and disease: norms from the georgia centenarian study*. *Exp Aging Res*, 2010. **36**(4): p. 394-425.
18. Mitchell, M.B., et al., *Norms from the Georgia Centenarian Study: measures of verbal abstract reasoning, fluency, memory, and motor function*. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn*, 2013. **20**(5): p. 620-37.
19. Rahman-Filipiak, A., et al., *Octogenarian and centenarian performance on the Fuld Object Memory Evaluation*. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn*, 2015. **22**(4): p. 438-51.
20. Holstege, H., et al., *The 100-plus Study of Dutch cognitively healthy centenarians: rationale, design and cohort description*. *bioRxiv*, 2018.
21. Folstein, M.F., S.E. Folstein, and P.R. McHugh, *"Mini-mental state". A practical method for grading the cognitive state of patients for the clinician*. *J Psychiatr Res*, 1975. **12**(3): p. 189-98.
22. Kok, R.M. and F.R.J. Verhey, *Gestandaardiseerde MMSE*. 2002, Zeist: Altrecht GGZ.
23. Nelson, H.E. and A. O'Connell, *Dementia: the estimation of premorbid intelligence levels using the New Adult Reading Test*. *Cortex*, 1978. **14**(2): p. 234-44.

24. Schmand, B., et al., [*The Dutch Reading Test for Adults: a measure of premorbid intelligence level*]. Tijdschr Gerontol Geriatr, 1991. **22**(1): p. 15-9.
25. Schmand, B., J. Lindeboom, and F. van Harskamp, *Dutch adult reading test*. 1992, Lisse, the Netherlands: Swets & Zeitlinger.
26. Wechsler, D., *WAIS-III: Wechsler adult intelligence scale*. 1997, San Antonio: TX: Psychological Corporation.
27. Benton, A.L., Hamsher, K. D., & Sivan, A., *Multilingual Aphasia Examination*. Neuropsychology, blocking, schizophrenia. Vol. 59. 1989, Iowa City, IA: AJA Associates.
28. Schmand, B., S.C. Groenink, and M. van den Dungen, [*Letter fluency: psychometric properties and Dutch normative data*]. Tijdschr Gerontol Geriatr, 2008. **39**(2): p. 64-76.
29. Lezak, M.D., *Neuropsychological assessment*. 2004, USA: Oxford University Press.
30. Van der Elst, W., et al., *Normative data for the Animal, Profession and Letter M Naming verbal fluency tests for Dutch speaking participants and the effects of age, education, and sex*. J Int Neuropsychol Soc, 2006. **12**(1): p. 80-9.
31. Balen, H.G.G., & Groot Zwaafink, A. J. M. , *The Rivermead Behavioural Memory Test. Nederlandstalige bewerking. Handleiding*. . 1987, Lisse: Swets & Zeitling.
32. Wilson, B., et al., *The development and validation of a test battery for detecting and monitoring everyday memory problems*. J Clin Exp Neuropsychol, 1989. **11**(6): p. 855-70.
33. Wilson, B.A., Cockburn, J., & Baddeley, A.D. , *The Rivermead Behavioural Memory Test*. . 1985, Bury St Edmunds, UK: Thames Valley Test Company.
34. Lindeboom, J. and C. Jonker, *Amsterdamse Dementie Screeningstest*. 1988, Lisse, the Netherlands: Swets and Zeitlinger.
35. Lindeboom, J., et al., *Visual association test to detect early dementia of the Alzheimer type*. J Neurol Neurosurg Psychiatry, 2002. **73**(2): p. 126-33.
36. Reitan, R.M., *Validity of the Trail Making Test as an indicator of organic brain damage*. Perceptual and Motor Skills, 1958. **8**(3): p. 271-276.
37. Wilson, B.A., et al., *Behavioural assessment of the dysexecutive syndrome*. 1996, St Edmunds, UK: Thames Valley Test Company.
38. Warrington, E.K. and M. James, *The visual object and space perception battery*. 1991, Bury St Edmunds: Thames Valley Test Company.
39. Shulman, K.I., *Clock-drawing: is it the ideal cognitive screening test?* Int J Geriatr Psychiatry, 2000. **15**(6): p. 548-61.
40. Shulman, K.I., et al., *Clock-drawing and dementia in the community: A longitudinal study*. International Journal of Geriatric Psychiatry, 1993. **8**(6): p. 487-496.
41. Munang, L., M. Chan, and W. Lim, *Diagnostic performance of the clock drawing test using a pre-drawn circle in persons with early dementia*. Asian Journal of Gerontology and Geriatrics, 2010. **5**: p. 54-61.
42. United Nations Educational, S.a.C.O. *International standard classification of education or ISCED-1997*. 1997; Available from: <http://www.unesco.org/education/>
43. Collin, C., et al., *The Barthel ADL Index: a reliability study*. Int Disabil Stud, 1988. **10**(2): p. 61-3.
44. de Haan, R., et al., *Clinimetric evaluation of the Barthel Index, a measure of limitations in daily activities*. Ned Tijdschr Geneesk, 1993. **May 1**(137 (18)): p. 917-21.
45. Mahoney, F.I. and D.W. Barthel, *Functional evaluation: the Barthel index*. 1965, Mary State Matiland GD.
46. Post, M.W.M., et al., *Nederlandse interviewversie van de Barthel-index onderzocht bij dwarslaesiepatiënten*. Nederlands Tijdschrift voor Geneeskunde, 1995. **139**: p. 1376-1376.
47. Yesavage, J.A. and J.I. Sheikh, *Geriatric Depression Scale (GDS) recent evidence and development of a shorter violence*. Clinical gerontologist, 1986. **5.1-2**: p. 165-173.

48. Kahle-Wroblewski, K., et al., *Sensitivity and specificity of the mini-mental state examination for identifying dementia in the oldest-old: the 90+ study*. J Am Geriatr Soc, 2007. **55**(2): p. 284-9.
49. Baldelli, M.V., et al., *A survey of a centenarian population in Italy, focusing on self-sufficiency and cognition*. Arch Gerontol Geriatr, 1996. **22 Suppl 1**: p. 345-54.
50. Bauco, C., Borriello, C., Cinti, A. M., Martella, S., Zannino, G., Rossetti, C., ... & Marigliano, V. , *Correlation Between MMSE performance, age and education in centenarians* Archives of Gerontology and Geriatrics, 1998. **26**: p. 23-26.
51. Mossakowska, M., et al., *Cognitive performance and functional status are the major factors predicting survival of centenarians in Poland*. J Gerontol A Biol Sci Med Sci, 2014. **69**(10): p. 1269-75.
52. Brumback-Peltz, C., et al., *Diagnosing dementia in the oldest-old*. Maturitas, 2011. **70**(2): p. 164-8.
53. Mitchell, A.J., *A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment*. Journal of Psychiatric Research, 2009. **43**(4): p. 411-431.
54. Tombaugh, T.N. and N.J. McIntyre, *The mini-mental state examination: a comprehensive review*. J Am Geriatr Soc, 1992. **40**(9): p. 922-35.
55. Black, S.A., et al., *The Influence of Noncognitive Factors on the Mini-Mental State Examination in Older Mexican-Americans*. Journal of Clinical Epidemiology, 1999. **52**(11): p. 1095-1102.
56. Marcopulos, B.A., C.A. McLain, and A.J. Giuliano, *Cognitive impairment or inadequate norms? A study of healthy, rural, older adults with limited education*. The Clinical Neuropsychologist, 1997. **11**(2): p. 111-131.
57. Sachdev, P.S., et al., *The Sydney Centenarian Study: methodology and profile of centenarians and near-centenarians*. Int Psychogeriatr, 2013. **25**(6): p. 993-1005.
58. Corrada, M.M., et al., *Prevalence of dementia after age 90: results from the 90+ study*. Neurology, 2008. **71**(5): p. 337-43.
59. Yang, Z., M.J. Slavin, and P.S. Sachdev, *Dementia in the oldest old*. Nat Rev Neurol, 2013. **9**(7): p. 382-93.

**Table 1. Demographic and clinical characteristics of the sample (n=235)**

Characteristic	
Age, median (IQR) <sup>a</sup>	100.4 (100.2-102)
Age range	(100-107)
Sex, N female (%)	168 (72)
Education, N (%)	
High level	90 (38)
Low level	145 (62)
Vision, N (%)	
Good	153(65)
Moderate	27 (12)
Poor	29 (12)
Very poor	22 (9)
Hearing, N (%)	
Good	134 (57)
Moderate	73 (31)
Poor	21 (9)
Very poor	4 (2)
Living situation, N (%)	
Independent without assistance, or in a residence with available service	138 (59)
In a residential care center	85 (36)
In a nursing home	2 (1)
With family	10 (4)
Barthel Index, N (%)	
≥15 Independent in ADL	126 (54)
<15 Dependent in ADL	80 (34)
GDS >5, depressive symptoms, N (%)	19 (8)
Mobility, N (%)	
Able to walk independently <sup>b</sup>	185 (79)
Able to walk with help of another person	8 (3)
Able to move independently in a wheelchair	14 (6)
Not able to move independently in a wheelchair	12 (5)

Values are presented by N and (%), or <sup>a</sup>median (interquartile range). ADL=Activities of Daily Living. GDS=Geriatric Depression Scale. There were 3 missing values for hearing, 4 for vision, 16 for mobility, 45 for the GDS and 29 for ADL independence. <sup>b</sup>=with or without help of a walking stick or walker. High education level=short-cycle tertiary or higher. Low education level=post-secondary non-tertiary or lower.

**Table 2.** Overview of sample sizes used for generating normative data

Tests	Tests presented, <i>N</i>	Tests completed, <i>N</i> (%)	Reasons for test incompleteness, %					Tests after exclusion for sensory losses, <i>N</i>	
			No compr.	Fatigue	Hearing	Vision	Several		Other <sup>C</sup>
MMSE	235	177 (75)	0	0	9	60	12	19	151 <sup>ab</sup>
RBMT	201	175 (87)	0	23	42	0	0	35	167 <sup>b</sup>
Number Location	201	152 (76)	16	10	0	63	0	10	142 <sup>a</sup>
Key Search	209	152 (73)	14	19	0	47	4	16	138 <sup>a</sup>
Clock Drawing Test	234	181 (77)	0	13	0	60	6	21	162 <sup>a</sup>
Letter Fluency	226	214 (95)	0	50	0	0	0	50	214
Animal Fluency 1 min	203	196 (97)	0	43	0	0	0	57	196
Animal Fluency 2 min	204	196 (96)	0	50	0	0	0	50	196
VAT Memory	229	178 (78)	0	22	0	57	0	22	156 <sup>a</sup>
VAT Naming	206	153 (74)	0	19	0	53	0	28	132 <sup>a</sup>
Digit Span Forward	218	178 (82)	0	40	13	0	3	45	163 <sup>b</sup>
Digit Span Backward	218	180 (83)	0	47	13	0	3	37	165 <sup>b</sup>
TMT A	202	133 (66)	1	26	0	49	7	16	127 <sup>a</sup>
TMT B	202	91 (45)	23	19	0	32	8	19	90 <sup>a</sup>
DART	228	169 (74)	0	24	0	58	2	17	153 <sup>a</sup>
Average	214	168 (79)	6	22	4	41	4	23	

Columns represent respectively: total number of tests that the centenarians were subjected to, total number of tests that could be completed, reasons for inability to complete tests and total number of tests used for generating normative data after exclusion for sensory losses. MMSE=Mini-Mental State Examination. RBMT=Rivermead Behavioural Memory Test. VAT=Visual Association Test. TMT=Trail Making Test. DART= Dutch Adult Reading Test. No compr.=No comprehension of tests and/or testinstructions. <sup>a</sup>=centenarians with poor to very poor vision were excluded. <sup>b</sup>=centenarians with poor to very poor hearing were excluded. <sup>c</sup>Other includes problems with test equipment, reasons were not reported, physical impairments (tremor or motor), or when there was no time left to finish the whole test battery.

**Table 3.** Percentiles and means for cognitive test scores for the total sample and stratified by educational attainment

Test	Group	N	M	SD	Percentiles							
					5th	10th	25th	50th	75th	90th	95th	
MMSE	Total	151	25.6	3.1	20	21	24	26	28	29	30	
	HE	51	26.5	3.0	20	22	25	27	29	30	30	
	LE	100	25.2	3.1	19	20	23	26	28	29	30	
RBMT Immediate Recall	Total	167	8.8	4.7	2	3.5	5	8	12	15	18	
	HE	64	10.0	5.2	2	4	6	9	12.5	18.5	21	
	LE	103	8.0	4.1	2	3	5	7	11	13.5	16	
RBMT Delayed Recall	Total	167	5.3	4.4	0	0	2	4.5	7.5	11.5	14	
	HE	64	6.3	5.0	0	1	2.5	6	9	13	16.5	
	LE	103	4.6	3.9	0	0	1.5	3.5	7	11	12.5	
Number Location	Total	142	8.5	2.0	4	5	8	9	10	10	10	
	HE	47	9.0	1.2	6	7	9	9	10	10	10	
	LE	95	8.2	2.2	4	4	7	9	10	10	10	
Key Search	Total	138	6.7	3.6	2	3	4	5	9	13	14	
	HE	44	8.4	3.3	4	4	6	9	11	13	15	
	LE	94	5.8	3.5	2	3	3	5	7	12	14	
Clock Drawing Test	Total	162	3.4	1.3	1	2	2	3	5	5	5	
	HE	55	3.8	1.3	1	2	3	5	5	5	5	
	LE	107	3.2	1.3	1	2	2	3	4	5	5	
Letter Fluency	Total	214	24.4	10.5	9	11	17	24	32	38	43	
	HE	79	29.2	10.8	10	16	21	28	35	45	47	
	LE	135	21.6	9.3	7	9	14	21	30	33	38	
Animal Fluency 1 min	Total	196	11.4	4.3	6	6	8	11	15	17	19	
	HE	71	12.2	4.7	5	6	8	12	16	19	20	
	LE	125	10.9	3.9	6	6	8	11	14	16	18	
Animal Fluency 2 min	Total	196	17.2	6.7	8	9	12	17	21	26	30	
	HE	70	18.6	7.9	7	8	12	18	24	30	33	
	LE	126	16.4	5.7	8	9	12	17	20	25	27	
VAT Memory	Total	156	9.0	3.3	2	4	7	10	12	12	12	
	HE	50	9.9	2.9	2	5	9	11	12	12	12	
	LE	106	8.6	3.4	2	4	6	10	12	12	12	
VAT Naming	Total	132	11.5	1.1	10	10	11	12	12	12	12	
	HE	43	11.7	0.6	11	11	12	12	12	12	12	
	LE	89	11.4	1.2	9	10	11	12	12	12	12	
Digit Span Forward score	Total	163	7.1	1.8	4	5	6	7	8	10	10	
	HE	64	8.0	1.8	4	6	7	8	9	10	11	
	LE	99	6.6	1.6	4	5	5	6	8	9	10	
Digit Span Forward span	Total	160	5.1	1.1	3	4	4	5	6	6	7	
	HE	63	5.5	1.0	3	4	5	6	6	7	7	
	LE	97	4.8	1.0	3	4	4	5	6	6	6	
Digit Span Backward score	Total	165	4.6	1.4	2	3	4	5	5	6	8	
	HE	64	5.0	1.4	3	3	4	5	6	7	8	
	LE	101	4.4	1.4	2	3	3	4	5	6	7	
Digit Span Backward span	Total	163	3.8	0.9	2	3	3	4	4	5	5	
	HE	64	3.9	0.7	3	3	3	4	4	5	5	
	LE	99	3.7	1.0	2	3	3	4	4	5	5	
TMT A	Total	127	113.1	66.9	258	199	134	92	70	58	51	
	HE	40	104.8	63.9	260	202	109	85	65	56	55	
	LE	87	116.9	68.3	266	199	140	98	70	60	49	
TMT B	Total	90	310.9	171.9	753	567	376	286	178	130	113	
	HE	34	267.0	154.6	591	436	299	258	162	122	92	
	LE	56	337.5	177.7	763	628	417	303	221	132	118	
DART IQ score	Total	153	98.4	13.9	75	79	87	99	108	118	122	
	HE	52	110.6	9.9	94	98	104	112	119	124	126	
	LE	101	92.1	11.3	74	78	84	92	101	106	110	

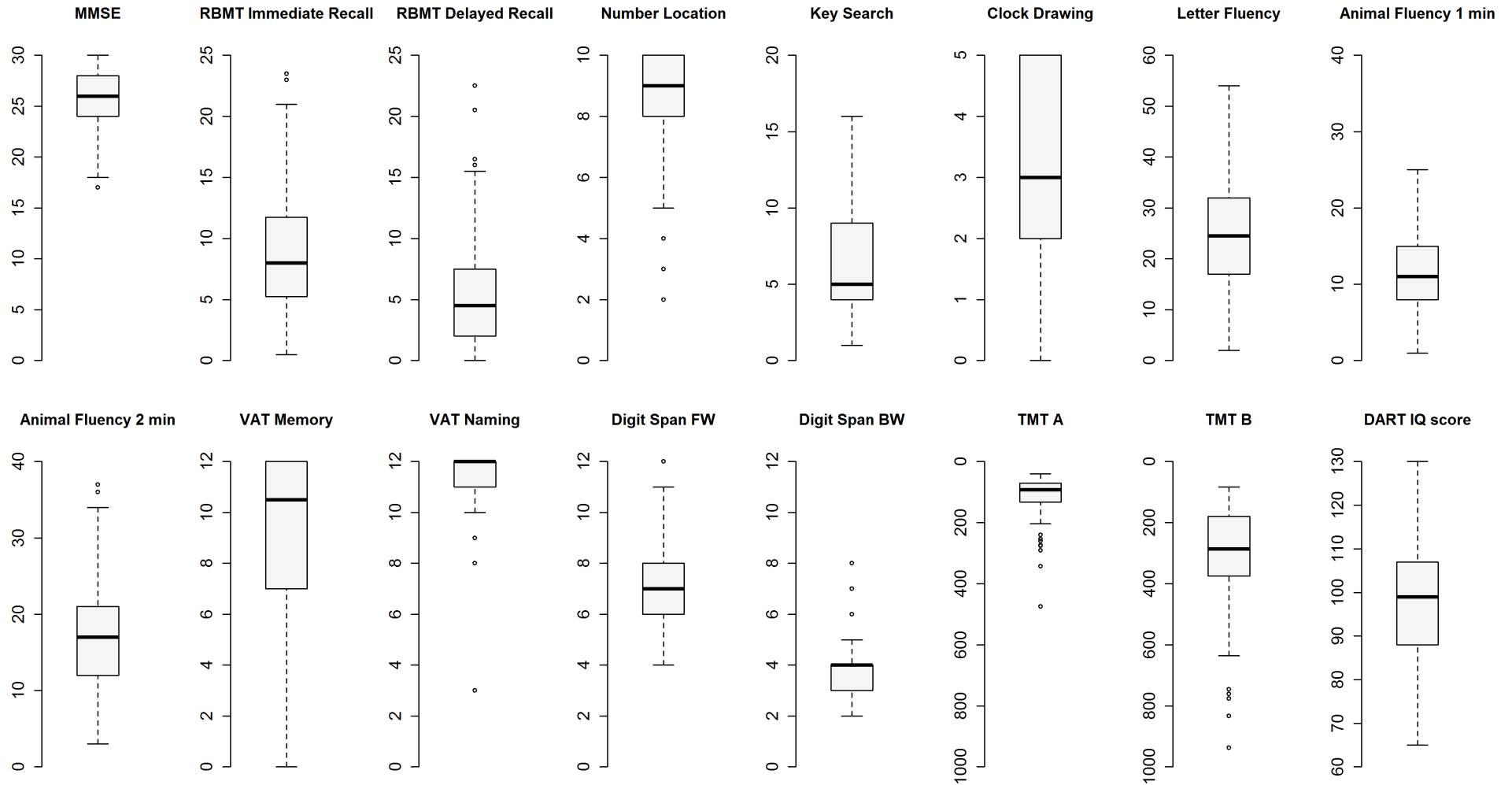
HE=High education (short-cycle tertiary or higher). LE=Low education (post-secondary non-tertiary or lower). MMSE=Mini-Mental State Examination. RBMT=Rivermead Behavioural Memory Test. VAT=Visual Association Test. TMT=Trail Making Test. DART= Dutch Adult Reading Test.

**Table 4.** Multiple linear regression analyses with sex, age and education as independent variables and cognitive test outcome as dependent variable

Tests	$R^2$	Sex			Age			Education		
		<i>B</i>	<i>SE</i>	<i>p</i>	<i>B</i>	<i>SE</i>	<i>p</i>	<i>B</i>	<i>SE</i>	<i>p</i>
MMSE	0.08	-0.60	0.58	0.30	-0.28	0.23	0.23	0.51	0.17	<b>&lt;.001</b>
RBMT Immediate Recall	0.05	-0.64	0.79	0.42	-0.10	0.31	0.73	0.61	0.24	<b>0.01</b>
RBMT Delayed Recall	0.03	-0.70	0.76	0.36	-0.10	0.29	0.73	0.45	0.23	0.05
Key Search	0.16	-1.71	0.65	<b>0.01</b>	-0.04	0.26	0.86	0.75	0.19	<b>&lt;.001</b>
Letter Fluency	0.16	3.74	1.50	<b>0.01</b>	-0.01	0.51	0.98	2.66	0.43	<b>&lt;.001</b>
Animal Fluency 1 min	0.05	1.24	0.67	0.06	-0.17	0.23	0.46	0.54	0.20	<b>0.01</b>
Animal Fluency 2 min	0.05	1.72	1.05	0.10	-0.26	0.36	0.47	0.94	0.31	<b>&lt;.001</b>
Digit Span Forward <i>score</i>	0.16	-0.59	0.29	<b>0.04</b>	0.09	0.12	0.43	0.41	0.09	<b>&lt;.001</b>
Digit Span Forward <i>span</i>	0.11	-0.15	0.18	0.40	0.03	0.07	0.64	0.23	0.05	<b>&lt;.001</b>
Digit Span Backward <i>score</i>	0.06	0.20	0.24	0.42	0.01	0.10	0.92	0.23	0.07	<b>&lt;.001</b>
Digit Span Backward <i>span</i>	0.04	0.12	0.16	0.44	-0.04	0.06	0.52	0.12	0.05	<b>0.01</b>
TMT A	0.07	0.05	0.09	0.63	-0.10	0.04	<b>0.01</b>	0.05	0.03	0.09
TMT B	0.07	0.23	0.12	0.06	-0.05	0.05	0.38	0.06	0.03	0.08

Values are presented as unstandardized Beta (*B*) and standard error (*SE*). MMSE=Mini-Mental State Examination. RBMT=Rivermead Behavioural Memory Test. VAT=Visual Association Test. TMT=Trail Making Test. DART= Dutch Adult Reading Test. TMT scores were log-transformed.





**Figure 1. Distribution of neuropsychological test scores.** Boxplots represent raw test scores. MMSE=Mini-Mental State Examination. RBMT=Rivermead Behavioural Memory Test. VAT=Visual Association Test. FW=Forward. BW=Backward. TMT=Trail Making Test. DART= Dutch Adult Reading Test.