

Swedish Normative Data for Mindmore: A Comprehensive Cognitive Screening Battery, Both Digital and Self-Administrated

Wobbie van den Hurk¹ , Ingvar Bergman¹, Alejandra Machado^{1,2}, Jonas Bjeremo³  and Anders Gustavsson^{4,5,*}

¹Mindmore AB, Stockholm, Sweden

²Division for Clinical Geriatrics, Centre for Alzheimer Research, Department of Neurobiology, Care Sciences and Society, Karolinska Institute, Stockholm, Sweden

³Department of Statistics, Stockholm University, Stockholm, Sweden

⁴Quantify Research, Stockholm, Sweden

⁵Division of Neurogeriatrics, Department of Neurobiology, Care Sciences and Society, Karolinska Institute, Stockholm, Sweden

(RECEIVED October 20, 2020; FINAL REVISION February 9, 2021; ACCEPTED March 12, 2021; FIRST PUBLISHED ONLINE May 24, 2021)

Abstract

Objective: Cognitive impairment is a key element in most mental disorders. Its objective assessment at initial patient contact in primary care can lead to better adjusted and timely care with personalised treatment and recovery. To enable this, we designed the Mindmore self-administrative cognitive screening battery. What is presented here is normative data for the Mindmore battery for the Swedish population. **Method:** A total of 720 healthy adults (17 to 93 years) completed the Mindmore screening battery, which consists of 14 individual tests across five cognitive domains: attention and processing speed, memory, language, visuospatial functions and executive functions. Regression-based normative data were established for 42 test result measures, investigating linear, non-linear and interaction effects between age, education and sex. **Results:** The test results were most affected by age and to a lesser extent by education and sex. All but one test displayed either linear or accelerated age-related decline, or a U-shaped association with age. All but two tests showed beneficial effects of education, either linear or subsiding after 12 years of educational attainment. Sex affected tests in the memory and executive domains. In three tests, an interaction between age and education revealed an increased benefit of education later in life. **Conclusion:** This study provides normative models for 14 traditional cognitive tests adapted for self-administration through a digital platform. The models will enable more accurate interpretation of test results, hopefully leading to improved clinical decision making and better care for patients with cognitive impairment.

Keywords: Neuropsychology, Primary health care, Cognitive ageing, Adults, Clinical decision-making, Telemedicine, Neuropsychological tests, Psychometrics

INTRODUCTION

Mental disorders are the leading cause of morbidity, affecting about a third of the European population (Wittchen et al., 2011). Their associated costs were estimated at €800 billion in 2010 in Europe alone, by far exceeding the total costs of cancer and heart disease (Gustavsson et al., 2011). The two most burdensome of these disorders are depression, with 264 million sufferers globally (James et al., 2018), and dementia which, due to the ageing of populations, grows by about 10 million people globally every year (WHO, 2019). Clearly, the management of these disorders should be the top priority to healthcare globally.

Cognitive impairment is a key element in most mental disorders, including neurodegenerative disorders (Aarsland et al., 2017; Carter et al., 2012; Grzegorski & Losy, 2017), mood disorders (Grossi et al., 2015; Hammar & Årdal, 2009) and psychiatric disorders (Menkes et al., 2019; Yehuda et al., 2006). Comprehensive cognitive screenings are, however, uncommon with suspected mental disorders (Iracleous et al., 2010; Socialstyrelsen, 2017, 2018) even though they can lead to a multitude of benefits. Cognitive testing can assist faster diagnosis (Woolley et al., 2011; Zucchella et al., 2018) and better-adjusted care (Withall et al., 2009; Zuckerman et al., 2018). In some cases, it can contribute to predicting progression to amnesic MCI (Gustavsson et al., 2020). Finally, repeated testing can support evaluation of treatment and recovery by providing an objective measurement (Levey et al., 2006; Schmid & Hammar, 2013; Solé et al., 2017).

*Correspondence and reprint requests to: Anders Gustavsson, H1 Neurobiologi, vårdvetenskap och samhälle, H1 Neurogeriatrik Winblad, 171 77 Stockholm, Sweden. Email: anders.gustavsson@ki.se

Primary care is the first point of contact for patients (WHO, 2018, p. 13). As such, it is advisable that one offers a first assessment of cognitive impairment already in this setting to identify patients for efficient management and referral and to relieve the burden on specialist care (Hlávka et al., 2018). According to primary care physicians, the cognitive screening instrument that could accomplish this should be quick and easy to administer but also be comprehensive and sensitive to detect subtle cognitive impairments (Iracleous et al., 2010).

As such, digitised cognitive tests can offer great improvements in the detection of cognitive impairment in primary care (Sabbagh, Boada, Borson, Chilukuri, et al., 2020). With patient self-administration and automated scoring of results, already validated tests can be used more efficiently, at a larger scale and relatively low cost (Bauer et al., 2012; Miller & Barr, 2017; Sabbagh, Boada, Borson, Doraiswamy, et al., 2020; Tierney & Lermer, 2010). Furthermore, the recent COVID-19 pandemic has highlighted the importance of digital healthcare (Keesara et al., 2020), not least in mental care since COVID-19 has created a new group of patients with suspected cognitive impairment (Troyer et al., 2020).

Mindmore is a comprehensive cognitive screening battery designed for primary care. The battery currently includes 14 traditional cognitive tests (Lezak et al., 2012), digitised and adapted for self-administration. Together these tests cover five cognitive domains: attention and processing speed, memory, language, visuospatial functions and executive functions. The concordance of the digital and analogue versions has been demonstrated in a previous study with 81 Swedish adults aged 21 to 85 years (Björngrim et al., 2019). The tests were administered in both versions, counterbalanced and with a 4-week interval. The test versions showed moderate to high correlations for all tests, statistical score equivalence for all but one of the tests (Lakens, 2017), and test results of both versions were similarly affected by age. However, when translating cognitive tests to a new medium some changes to the results are unavoidable (Bauer et al., 2012; Germine et al., 2019). Thus, Mindmore-specific normative data were deemed necessary to aid clinicians in interpreting their patients' test results.

The current study aims to provide regression-based normative data for the adult Swedish population for each of the tests included in Mindmore. Data were collected in a controlled laboratory environment. The method of regression-based normative data is much preferred over traditional discrete normative data because it solves the problems of arbitrary age bands, varying sample sizes per normative group (Knight et al., 2006), and substantially decreases the amount of data required for accurate norms (Oosterhuis et al., 2016). The hypothesised influence on test results of the demographic variables age (Cohen et al., 2019), educational attainment (Ritchie et al., 2015) and sex (Miller & Halpern, 2014) were investigated.

METHODS

Study Design

A cross-sectional normative study was conducted, administering Mindmore to healthy volunteers. The study ran from June 2018 to February 2020. Testing took place in Stockholm and Uppsala, Sweden. All participants were previously unfamiliar with Mindmore and were unaware of the tasks that would be asked of them in the tests. Testing included an interview on current health status and medical history, a short cognitive screening test, and the Mindmore screening battery. Participants did not receive any financial remuneration for participating but were provided with a summary of their results at the end of the study. All participants received written information on the study prior to testing and signed informed consent. The study was approved by the Regional Ethics Review Board in Stockholm (issue number 2019-02030).

Participants

Individuals living in the community aged 17 years and above were invited to the study. The study was advertised on a specialized website for research participant recruitment (*StudentKaninen.se*, 2020), in targeted social media posts, with posters on public message boards, during pensioners' association meetings and by word of mouth. Recruitment was monitored to represent the full adult Swedish population (Statistiska centralbyrån, 2018a, 2018b) in terms of age, education and sex.

Exclusion criteria for the present study consisted of (1) a non-fluent level of Swedish; (2) previous experience of any of the Mindmore tests and; (3) impaired cognitive functioning as assessed by one of two short cognitive screening tests: three or more errors on the Short Portable Mental State Questionnaire (SPMSQ; Pfeiffer, 1975; Svedjebrant & Lindmark, 2011) or a score of below 25 on the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975); (4) an active central nervous system disease or psychiatric condition and; (5) prior history of disorders that could potentially affect cognition (e.g., evident stroke, significant head trauma, recent or current substance abuse). Treated chronic medical problems (e.g., diabetes, hypertension, cardiac problems) were not an exclusion criterion. Each participant completed a medical interview (Supplementary Figure S1). The test leader went over the interview answers, requesting further details (e.g. duration, type and dose of medication).

From the initial sample of 762 participants, 37 participants were excluded after the medical interview, three were excluded because they made too many errors on the SPMSQ screening test and two were excluded due to technical issues. The final sample included 720 participants (414 women and 306 men) and had an age range from 17 to 93 with a mean of 50.8 (SD 18.8). Their level of education ranged from 7 to 26 years of schooling with a mean of 15.1 (SD 2.9) years (Table 1). The results of the short cognitive screening tests,

Table 1. Demographics of the final sample

Ages	<i>n</i>	Age		Education		Sex	
		<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	Female	Male
<20	33	17.82	.81	10.94	.93	18	15
20 – 29	88	24.61	3.07	14.55	1.92	40	48
30 – 39	107	33.92	2.93	16.36	2.66	43	64
40 – 49	99	44.62	2.88	15.96	2.92	51	48
50 – 59	128	54.76	2.74	15.34	2.68	79	49
60 – 69	121	64.80	3.01	15.02	2.79	84	37
70 – 79	118	73.84	2.52	15.25	3.23	79	39
80+	26	84.08	4.18	13.54	3.08	20	6
Total	720	50.76	18.78	15.14	2.93	414	306

which were administered to 97% of the final sample, were as follows: the SPMSQ ($n = 652$) ranged from 0 to 2 with a mean of .6 ($SD .7$); and the results of MMSE ($n = 47$) ranged from 27 to 30 with a mean of 28.9 ($SD .9$). The remaining 21 participants, aged 17 or 18 years and full-time high school students, were exempt from the short cognitive screening tests. Participants were all from the Swedish population (93% were Swedish natives, the remainder was at least fluent in Swedish).

Battery Design

The Mindmore screening battery was designed to provide a comprehensive measure of cognitive functioning for the assessment of any patient with suspected cognitive impairment. The available literature on cognitive testing varies in terms of considered domains and how they are grouped (Lezak et al., 2012; Sachdev et al., 2014). We identified five domains which in our experience are common in clinical neuropsychology practice and supported by renowned international institutions and expert groups (American Psychiatric Association, 1994; Grzegorski & Losy, 2017; Nasreddine et al., 2005). These constitute attention and processing speed, memory, language, visuospatial functions and executive functions.

A total of 14 digitised traditional tests are included in the Mindmore application. These are tests clinically accepted in Europe (Maruta et al., 2011), with proven validity and reliability (Lezak et al., 2012), and possible to digitise while retaining correspondence with the traditional test instructions and stimuli (Björngrim et al., 2019).

Attention and processing speed were assessed using the Trail Making Test (TMT; Army Individual Test battery, 1944) part A to measure mental speed; the Symbol Digit Modalities Test (SDMT; Smith, 1982) to measure visual scanning; and an in-house test of simple reaction time (Simple RT Test; Mindmore screening battery, 2020).

Memory was assessed using a 15-word auditory learning test (Rey Auditory Verbal Learning Test; RAVLT; Rey, 1964; Schmidt, 1996) with five learning trials, short-term recall of the 15 words after a distraction list, long-term recall

of the 15 words after 20 minutes and finally recognition of the 15 words among 15 distractors; a 10-word auditory and visual learning test (CERAD Word List Learning Test; Morris et al., 1989) with three learning trials, recall of the 10 words after five minutes and finally recognition of the 10 words among 10 distractors; and a spatial memory test (Corsi, 1972; Milner, 1971).

Language was assessed using the Token Test (Boller & Vignolo, 1966; De Renzi & Vignolo, 1962) to measure auditory comprehension; the Boston Naming Test (BNT) – 15 words (Jørgensen et al., 2017; Kaplan et al., 1983); and the FAS Word Fluency Test (Benton et al., 1989).

Visuospatial functions were assessed using the Cube Drawing Test (Necker, 1832); and the Clock Drawing Test (Critchley, 1953; Shulman et al., 1986).

Executive functions were assessed using the TMT part B to measure mental flexibility; an in-house test of complex reaction time (Complex RT Test; Mindmore screening battery, 2020), the Paced Auditory Serial Addition Test (PASAT; Gronwall, 1977; Gronwall & Sampson, 1974) to measure information processing; a silent Stroop Test (Strauss, Sherman, & Spreen, 2006; Stroop, 1935) to measure concentration effectiveness; and the Tower of Hanoi Test (Lucas, 1883) to measure the ability to plan ahead.

Where possible, performance index scores were calculated taking both accuracy and speed into account, so called “throughput measures” (Thorne, 2006). Such index scores were calculated for the following tests: TMT part A and B [correct connections/time to completion], Simple RT Test [correct responses/reaction time], Complex RT Test [(correct responses – inhibition errors)/reaction time], Stroop Test impulse control [correct responses CW/reaction time CW], where CW stands for conflicting (and W for non-conflicting) colour of the ink and the word, and Tower of Hanoi [correct solution/time to completion]. In addition, index scores were calculated for RAVLT recognition [hits – false alarms], CERAD recognition [hits – false alarms] and for Stroop Test inhibition [reaction time CW – reaction time W]. The 14 tests in the Mindmore battery resulted in 42 test result measures, including the index scores and detailed scores on accuracy and speed.

Procedure

Participants were administered the Mindmore battery via a digital touch screen tablet of 12.3 inches. The battery consisted of a selection of the 14 tests. Participants who indicated they were colour blind were exempt from the Stroop Test, Simple and Complex RT Tests and the Token Test. The RAVLT, CERAD, BNT and FAS Test were only administered to native Swedish speakers. Participants over 75 years old were exempt from the Simple- and Complex RT Tests, PASAT and Tower of Hanoi. Not every test was available from the start of the study; RAVLT and CERAD cannot be included in the same assessment.

Scoring was fully automated except for CERAD, FAS Test, Cube Drawing Test and Clock Drawing Test. Speech data on RAVLT, CERAD, BNT and FAS Test were processed through Google speech recognition software (*Google Cloud Speech-to-Text*, 2020). The automated corrections were validated manually for CERAD and FAS Test by comparing the speech recordings to the automated results. The Cube and Clock Drawing Tests resulted in images that were manually scored similarly to common clinical practice (e.g. Nasreddine et al., 2005). That is, as 0 = not correct and 1 = correct, where .5 points each were given for correct numbers and correct hands on the Clock.

Testing took place in a quiet room free of distractions. A trained test leader conducted the medical interview and short screening test (MMSE or SPMSQ) with the participant after which the participant was left alone in the room to complete the Mindmore screening battery (approx. 45 min). Participants who indicated that they were unfamiliar with the use of a touch screen were given the option of supervised testing in which the test leader stayed in the room and gave minimal advice on tablet use during instructions and practice trials. The test protocol ran automatically, allowing the participant to continue to the next test when the previous test was finished. Each test consisted of an introduction both in audio and visualized on-screen, when applicable an interactive practice trial, and the test trial. Instructions could be repeated on demand. The participant received feedback during practice trials and, with multiple unsuccessful practice attempts, the specific test was cancelled and the test protocol continued to the next test.

Data Analysis

Multiple linear regression models were fitted on the data. The forward-backward stepwise predictor selection was employed with the Bayesian information criterion (BIC) selection criterion, as proposed by Chowdhury and Turin (2020). The initial models contained all candidate lower-order predictors to be primarily considered for exclusion, the candidate higher-order predictors were outside the model and primarily considered for inclusion.

The candidate lower-order predictors were age, years of education and sex. The candidate higher-order predictors were the quadratic term for age and two theory-driven interactions. The interaction age x years of education was included to capture potential effects of cognitive reserve (Stern, 2002). The theoretical concept of cognitive reserve accounts for the discrepancy, correlated to the level of education, between detected brain deterioration and task performance observed in some individuals and not in others (Stern & Barulli, 2019). The second interaction, age x sex, was included to capture a potential specific younger male advantage previously observed in digital reaction time-based tests (Feenstra et al., 2018). Lower-order predictors were forced back into the model when included in a higher-order predictor of the final model.

Furthermore, an alternative model was built replacing years of education with its inverse (1/years of education). The model with the inverse term was retained if it decreased the BIC value and if visual inspection of the model confirmed the improved curve fit. Otherwise, the model with years of education was selected as the final model (Bergman et al., 2016).

The assumptions of regression analysis were tested for each model. The occurrence of multicollinearity between lower-order predictors was checked by calculating the Variance Inflation Factors (VIFs), which should not exceed 10 (Yoo et al., 2014). Possible outliers and unusual observations with high leverage were identified by calculating Cook's distances, with the criterion of $D_i > .85$ (McDonald, 2002). The normality of the residuals' distribution was evaluated by visual inspection and by applying the Kolmogorov-Smirnov test. The homoscedasticity of the residuals' distribution was evaluated by visual inspection and by applying the Levene's test on the residuals divided into quartiles of predicted score (Testa et al., 2009; Van der Elst et al., 2006). When homoscedasticity could not be assumed, the SD of residuals was calculated for each quartile. An alpha level of 1% was applied due to the increased likelihood of obtaining statistical significance in a sample of this size. All analyses were performed in R version 3.6.0 (R Core Team, 2019).

Normative Model Application

Normative regression models and variance of the model residuals can be used to calculate the standardised normative value for test takers. As such, clinicians receive a standardised measure of deviation for their patient's observed test result compared to their expected normative test result (their predicted score). The two formulae needed to calculate the standardised normative value (z) are depicted below. First, the predicted score for each test result is calculated:

$$\begin{aligned} \text{predicted score} = & \text{Intercept} + (B_{\text{age}} \times \text{age}) + (B_{\text{age}^2} \times \text{age}^2) \\ & + [(B_{\text{education}} \times \text{education}) \text{ OR } (B_{1/\text{education}} / \text{education})] \\ & + (B_{\text{age} \times \text{education}} \times \text{age} \times \text{education}) \\ & + (B_{\text{sex}} \times \text{sex}) + (B_{\text{age} \times \text{sex}} \times \text{age} \times \text{sex}) \end{aligned}$$

Second, the difference between the observed and predicted scores of the test taker is standardised:

$$z\text{-score} = (\text{observed score} - \text{predicted score}) / \text{SD of residuals}$$

For a more detailed explanation see Van der Elst et al. (2006). For the details necessary to employ an alternative method devised by Crawford and Howell (1998) see the Supplementary Methods.

Table 2. Descriptive summary of results from the self-administered Mindmore cognitive tests

Cognitive domains and test	Direction of impairment (logical range)	<i>n</i>	<i>Mean</i>	<i>SD</i>	Observed range	
					min	max
Attention and processing speed						
TMT part A index	↘ NA	550	.86	.30	.20	2.30
Time to completion (s)	↗ (max. 120)	550	31.92	14.47	10.43	120.00
Correct connections	↘ (24 to 0)	550	23.83	.61	19	24
SDMT	↘ NA	199	42.44	13.44	0	79
Simple RT Test index	↘ NA	416	18.29	3.31	8.54	28.69
Reaction time (ms)	↗ NA	416	502.98	93.84	313.67	871.44
Correct responses	↘ (9 to 0)	416	8.91	.39	6	9
Memory						
RAVLT learning	↘ (75 to 0)	372	50.36	9.19	23	70
RAVLT short-term recall	↘ (15 to 0)	439	10.73	2.97	1	15
RAVLT long-term recall	↘ (15 to 0)	437	10.60	3.11	2	15
RAVLT recognition index	↘ (15 to -15)	450	12.98	2.49	1	15
Hits	↘ (15 to 0)	450	14.08	1.35	7	15
False alarms	↗ (0 to 15)	450	1.11	1.87	0	13
CERAD learning	↘ (30 to 0)	186	21.44	3.90	11	30
CERAD recall	↘ (10 to 0)	186	7.17	2.18	0	10
CERAD recognition index	↘ (10 to -10)	212	9.55	.98	5	10
Hits	↘ (10 to 0)	212	9.73	.72	6	10
False alarms	↗ (0 to 10)	212	.17	.49	0	4
Corsi Span forward	↘ (9 to 2)	692	6.13	1.18	2	9
Span backward	↘ (9 to 2)	680	5.59	1.21	2	9
Language						
Token Test	↘ (36 to 0)	183	33.04	3.15	19	36
BNT	↘ (15 to 0)	536	11.99	2.59	0	15
FAS Test	↘ NA	184	45.69	13.86	9	85
Visuospatial functions						
Cube Drawing Test	↘ (1 to 0)	467	.88	.33	0	1
Clock Drawing Test	↘ (1 to 0)	164	.86	.30	0	1
Executive functions						
TMT part B index	↘ NA	554	.36	.17	.00	.91
Time to completion (s)	↗ (max. 240)	554	73.82	37.45	17.73	240.00
Correct connections	↘ (24 to 0)	554	21.92	4.31	0	24
Complex RT Test index	↘ NA	395	15.22	3.43	2.39	24.47
Reaction time (ms)	↗ NA	395	661.46	140.16	388.50	1249.60
Correct responses	↘ (10 to 0)	395	9.86	.53	3	10
Inhibition errors	↗ (0 to 33)	395	.23	.53	0	4
PASAT	↘ (60 to 0)	256	48.00	9.99	8	60
Stroop Test index	↘ NA	552	12.46	3.46	.38	24.26
Inhibition (ms)	↗ NA	552	395.92	387.31	-1146.52	3779.70
Reaction time W (ms)	↗ NA	552	1307.61	431.21	743.25	3696.39
Correct responses W	↘ (20 to 0)	552	19.86	.44	15	20
Reaction time CW (ms)	↗ NA	552	1703.53	543.22	824.50	5255.00
Correct responses CW	↘ (20 to 0)	552	19.52	1.45	2	20
Tower of Hanoi index	↘ NA	209	.73	.46	.00	2.22
Correct solution	↘ (1 to 0)	209	.98	.14	0	1
Time to completion (min)	↗ NA	205	2.10	1.82	.45	12.00

TMT = Trail Making Test; SDMT = Symbol Digit Modalities Test; RT = reaction time; RAVLT = Rey Auditory Verbal Learning Test; CERAD = Word List Learning Test; BNT = Boston Naming Test; FAS = Word Fluency Test; PASAT = Paced Auditory Serial Addition Test; Stroop Test W = congruent word trial; Stroop Test CW = incongruent colourword trial; Logical range = theoretical min. and max. obtainable test result.

RESULTS

Data for 720 participants were entered into the analysis. The number of participants per test ranged from 164 for the Clock Drawing Test to 692 for the Corsi Span forward with a

median of 427 participants per test. A descriptive summary of the results is presented in Table 2 with a more detailed description in Supplementary Table S2. There was no potentially harmful multicollinearity among the lower-order predictors ($VIF \leq 1.164$). Some extreme results (>3 SD

deviation in either direction from the estimated regression line) were present in the data (mean 5.88 per test result, range 0 to 16). No individual had extreme results in more than two tests. The extreme results did not significantly influence any of the models (Cook's distance $\leq .294$) and might very well be part of normal population performance (Binder et al., 2009). Out of the 42 test residuals' distributions, 20 were not considered to deviate too much from a normal distribution with an absolute skew mean of .58 (range .02 to 1.87) and a mean kurtosis of .99 (range $-.33$ to 4.91). Three test result measures (Cube Drawing Test, Clock Drawing Test and Tower of Hanoi correct solution) have only two or three possible outcomes, with a high percentage of participants receiving the maximum result. As such, the residuals' distributions for these test result measures did not show enough continuous variance to be considered normally distributed. The residuals' distributions of the 19 remaining test result were either somewhat skewed, somewhat leptokurtic or both (absolute mean skew 3.34, range 1.53 to 7.64; mean kurtosis 19.51, range 4.84 to 79.61). These 19 test results were the Token Test (skew = -2.0 ; kurtosis = 4.8), six time measures and 12 measures of correct or incorrect responses.

Normative data were established for 42 test result measures (Table 3). The regression analyses yielded models for 33 of them including one or more demographic predictors. The remaining nine models which were unaffected by the demographic variables were all measures of accuracy of performance. Exemplary models illustrating the observed demographic effects are presented in Figures 1–3. Remaining models with corresponding effects are presented in Supplementary Figures S2 to S6.

Overall, age was the strongest predictor included in all models including predictors except for the FAS Test. A linear decrease in performance with increasing age was observed in 17 models. A curved age effect indicating accelerated decrement with age was seen in 15 models. TMT part A time to completion and SDMT showed accelerated slowing of performance from ages 30 and 20, respectively (Figure 1a). Accelerated performance decrement with older age was also seen for Corsi Span forward and backward from ages 25 and 40, respectively, and for the Cube Drawing Test. The latter also showed an interaction with education implying an accelerated performance decrement from age 30 for lower education but not before age 50 for higher education. In addition, slowly accelerating performance decline was seen for TMT part B index already from age 20 to 25, and for the Stroop Test reaction time W and CW already from age 30. Reaction time CW also showed an interaction with education with somewhat slower decline with age at higher education (Figure 1b).

Four models displayed a U-shaped age effect, in which performance increased up to age 50–60 after which an accelerated performance decrease took over. These were the Token Test and BNT in the language domain, the Clock Drawing Test and the Complex RT Test inhibition errors (Figure 1c). BNT also showed an interaction with education implying somewhat later onset of decline at higher education (Figure 1d).

Education was included in 24 models across every domain. An increase in number of years of education was consistently associated with higher performance. In 15 models, performance increased linearly with every additional year of education. In the other nine models, constituting the verbal memory tests, TMT part B and the PASAT Test, the effect was asymptotic (Figure 2). Here the most profound performance increase was associated with the first 12 years of education. Further years of education continued to add a small benefit, though the increments became smaller with every additional year.

Sex was included in 11 models. A female advantage was seen for RAVLT, CERAD recall and Stroop Test reaction time W (Figure 3a). A male advantage was observed for Corsi Span and Complex RT Test (Figure 3b). The interaction between age and sex was not included in any of the models.

In addition, Table 3 shows that the “throughput” index scores in most cases add to the variance explained as compared to the separate variables included in the indices. The only exception is Complex RT, where fewer inhibition errors seem to compensate for slower reaction times.

The variance in the test results, which remained after controlling for the demographic variables age, education and sex, was quantified through the residuals' distribution of each model (Table 4). For 16 out of 33 test result measures homoscedasticity of the residuals' distribution could not be assumed based on visual inspections and Levene's test ($p < .01$). For these test result measures, the SD of residuals was calculated for each quartile of the predicted score.

DISCUSSION

Mindmore is a screening battery of 14 cognitive tests, adapted for digital self-administration. Research over many decades has assessed the validity and reliability of the underlying traditional analogue tests (Lezak et al., 2012). Digitalisation of traditional tests can increase their availability and ease of administration (Bauer et al., 2012; Miller & Barr, 2017). Moreover, digitised tests potentially facilitate enhanced standardisation of assessment and result calculation (Newman et al., 2018). An equivalence study between the digitised Mindmore tests and the underlying traditional analogue tests showed comparable results between the two (Björngrim et al., 2019). However, complete correspondence of test results between analogue and digital test versions cannot be expected. Therefore, norms for the 14 Mindmore tests are provided in the present study. These norms are intended for the adult Swedish population and adjusted for age, education and sex.

The effects of the demographic variables on the test results collected with Mindmore largely resembled the effects previously reported for the traditional analogue tests (Lezak et al., 2012; Mitrushina et al., 2005; Strauss et al., 2006). Age demonstrated the strongest impact on test performance. Irrespective of cognitive domain, all but one of the final models including predictors declined with increasing age. These

Table 3. Final multiple linear regression models

Cognitive test	Intercept			Age			Age ²			Years of education			Age * years of education			Sex			R ²
	B	(SE)	p	B	(SE)	p	B	(SE)	p	B	(SE)	p	B	(SE)	p	B	(SE)	p	
TMT part A index	1.027	(0.061)	***	-0.008	(0.001)	***				0.017	(0.004)	***							0.298
Time to completion (s)	38.072	(4.064)	***	-0.337	(0.168)	*	0.007	(0.002)	***	-0.584	(0.189)	**							0.238
Correct connections	23.833	(0.026)	***																
SDMT	56.046	(3.987)	***	-0.062	(0.153)		-0.005	(0.001)	***	0.554	(0.169)	**							0.737
Simple RT Test index	19.900	(0.803)	***	-0.089	(0.009)	***				0.165	(0.051)	**							0.204
Time (ms)	457.631	(22.963)	***	2.453	(0.248)	***				-4.484	(1.460)	**							0.193
Correct responses	8.906	(0.019)	***																
RAVLT learning ^a	75.009	(2.493)	***	-0.252	(0.023)	***				-162.248	(30.035)	***				-3.195	(0.836)	***	0.286
RAVLT short-term recall ^a	17.265	(0.812)	***	-0.061	(0.007)	***				-44.509	(9.571)	***				-1.257	(0.267)	***	0.175
RAVLT long-term recall ^a	17.877	(0.853)	***	-0.064	(0.008)	***				-55.108	(10.123)	***				-1.018	(0.280)	***	0.180
RAVLT recognition index ^a	17.496	(0.685)	***	-0.051	(0.006)	***				-25.929	(7.873)	**				-0.672	(0.225)	**	0.145
Hits ^a	15.993	(0.386)	***	-0.018	(0.003)	***				-12.117	(4.438)	**				-0.468	(0.127)	***	0.081
False alarms	1.135	(0.587)		-0.043	(0.027)		0.001	(0.000)	**										0.104
CERAD learning ^a	31.794	(1.264)	***	-0.088	(0.012)	***				-75.973	(15.394)	***							0.307
CERAD recall ^a	12.790	(0.805)	***	-0.050	(0.008)	***				-33.755	(9.018)	***				-0.811	(0.303)	**	0.246
CERAD recognition index	9.552	(0.068)	***																
Hits	9.726	(0.050)	***																
False alarms	0.175	(0.034)	***																
Corsi																			
Span forward	5.829	(0.298)	***	0.017	(0.012)		-0.0005	(0.000)	***	0.044	(0.014)	**				0.273	(0.080)	***	0.271
Span backward	5.556	(0.295)	***	0.027	(0.013)	*	-0.001	(0.000)	***							0.368	(0.086)	***	0.198
Token Test	27.566	(1.637)	***	0.260	(0.066)	***	-0.003	(0.001)	***										0.091
BNT	8.452	(1.836)	***	0.218	(0.043)	***	-0.003	(0.000)	***	-0.281	(0.119)	*	0.006	(0.002)	**				0.160
FAS Test	31.939	(5.028)	***							0.921	(0.330)	**							0.041
Cube Drawing Test	1.273	(0.275)	***	-0.003	(0.006)		-0.0001	(0.000)	**	-0.031	(0.017)		0.001	(0.000)	**				0.115
Clock Drawing Test	0.735	(0.190)	***	0.022	(0.007)	**	-0.0002	(0.000)	***	-0.019	(0.007)	*							0.110
TMT part B index ^a	0.578	(0.062)	***	0.002	(0.002)		-0.0001	(0.000)	**	-1.852	(0.439)	***							0.255
Time to completion (s)	105.729	(10.663)	***	-1.392	(0.441)	**	0.021	(0.004)	***	-1.622	(0.494)	**							0.217
Correct connections	19.393	(1.330)	***	0.087	(0.055)	**	-0.001	(0.001)	*	0.147	(0.062)	*							0.081
Complex RT Test index	16.988	(0.838)	***	-0.097	(0.009)	***				0.144	(0.052)	**				0.994	(0.302)	**	0.261
Reaction time (ms)	496.300	(19.230)	***	4.063	(0.358)	***										-43.317	(12.077)	***	0.287
Correct responses	9.856	(0.027)	***																
Inhibition errors	1.102	(0.184)	***	-0.039	(0.009)	***	0.0004	(0.000)	***										0.062
PASAT ^a	64.696	(3.800)	***	-0.183	(0.037)	***				-128.073	(44.403)	**							0.099
Stroop Test index	18.268	(0.351)	***	-0.117	(0.007)	***													0.360
Inhibition (ms)	203.413	(48.241)	***	3.872	(0.914)	***													0.032
Reaction time W (ms)	1296.437	(130.928)	***	-8.086	(5.024)		0.2195	(0.049)	***	-15.512	(5.445)	**				97.036	(30.649)	**	0.353
Correct responses W	19.855	(0.019)	***																
Reaction time CW (ms)	911.547	(320.585)	**	4.936	(7.716)	**	0.2537	(0.064)	***	33.344	(21.684)		-0.879	(0.390)	*				0.347
Correct responses CW	19.524	(0.062)	***																
Tower of Hanoi index	1.262	(0.090)	***	-0.011	(0.002)	***													0.158
Correct solution	0.981	(0.010)	***																
Time to completion (min)	0.764	(0.379)	*	0.029	(0.008)	***													0.064

TMT = Trail Making Test; SDMT = Symbol Digit Modlities Test; RT = Reaction Time; RAVLT = Rey Auditory Verbal Learning Test; CERAD = Word List Learning Test; BNT = Boston Naming Test; FAS = Word Fluency Test; PASAT = Paced Auditory Serial Addition Test; Stroop Test W = congruent word trial; Stroop Test CW = incongruent colourword trial. Predicted score = Intercept + (B_{age} x age) + (B_{age²} x age²) + [(B_{education} x education) OR (B_{1/education} / education)] + (B_{age x education} x age x education) + (B_{sex} x sex). Where sex is 0 for women and 1 for men.

^a predictor years of education has been replaced with 1/years of education.

* p < .05; ** p < .01; *** p < .001

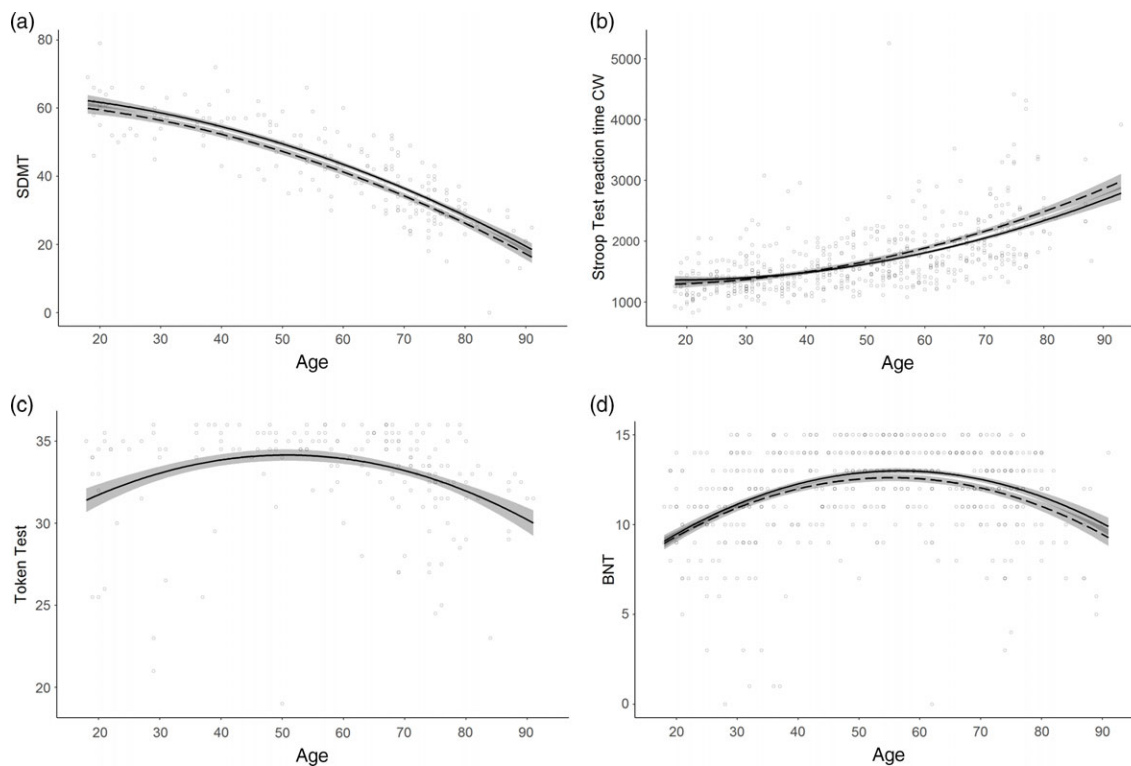


Fig. 1. Normative models with non-linear age effects. a. SDMT correct matches (Attention and processing speed) displaying a curved age effect; b. Stroop Test reaction time CW (milliseconds; Executive functions) displaying a curved age effect and an interaction between age and education; c. Token Test (Language) displaying a U-shaped age effect; d. BNT (Language) displaying a U-shaped age effect and an interaction between age and education. Line solidity indicates years of education: dashed = 12 years; solid (when paired with dashed) = 16 years; shaded area = confidence interval.

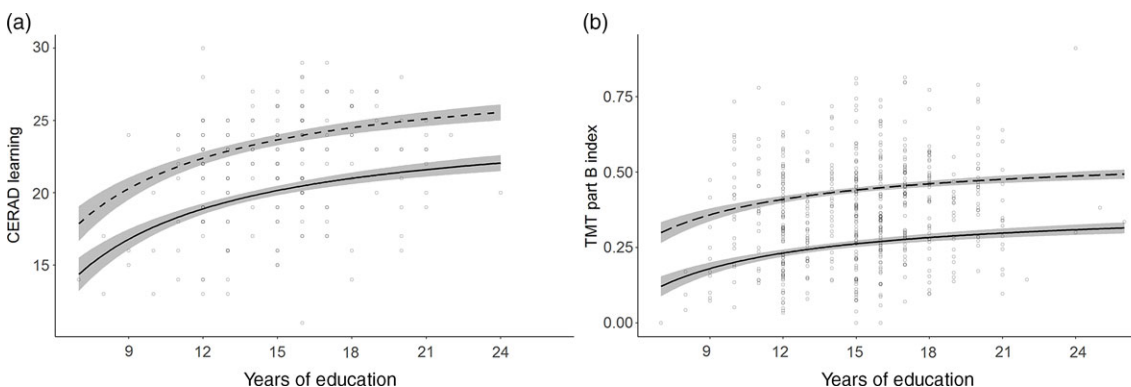


Fig. 2. Normative models with asymptotic education effects. a. CERAD Learning (Memory); b. TMT part B index (Executive functions) calculated as [correct connections/time to completion (s)]. Line solidity indicates age: dashed = 35 years; solid = 75 years; shaded area = confidence interval.

findings support previous findings in that age-related decline is known to occur in every cognitive domain, even though the onset and severity of decline varies between domains (Cohen et al., 2019). They also corroborate with previous findings maintaining that in general, a gradual and not necessarily linear decline in processing speed is observed from early middle age (Gautam et al., 2011; Zimprich & Mascherek, 2010). While basic attentional functions such as warning systems are preserved in normal ageing, more complex processes are not (Andrés et al.,

2006; McDowd & Shaw, 2000). The marked decline in processing speed is by some considered to be the first symptom of age-related cognitive decline (Finkel et al., 2003; Meijer et al., 2009; Soederberg Miller & Lachman, 2000). On the other end of the spectrum, there is consensus that functions in the language domain are relatively well retained in healthy elderly (Wingfield, 2000). This is confirmed by the results on the language tests in this study. An age-related decline was not seen until age 50–60 for BNT and Token Test, and not at all for the FAS Test.

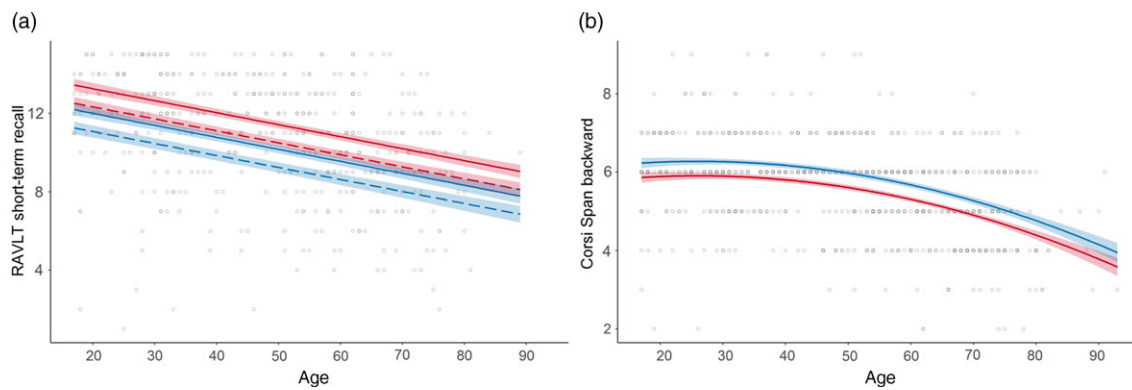


Fig. 3. Normative models with sex effects in the memory domain.

a. RAVLT short-term recall (Memory) displaying a female advantage; b. Corsi Span backwards (Memory) displaying a male advantage. Line colour indicates sex: red = female; blue = male; line solidity indicates years of education: dashed = 12 years; solid (when paired with dashed) = 16 years; shaded area = confidence interval.

Higher educational attainment was associated with improved performance on two-thirds of the test result measures, across all five cognitive domains. This finding is supported by large cohort studies reporting a benefit of education across a wide range of cognitive functions (Ritchie et al., 2015; Ritchie & Tucker-Drob, 2018). The benefit of education was either gradual or most profound for the early years of education. This latter pattern of a stronger effect of the earliest years of education was also found in a large study with older adults in the United States (Zahodne et al., 2015). In this study, the strongest effect of education is seen for the first 12 years of education. It might be speculated that this finding is related to the design of the Swedish educational system in which pupils graduate from their secondary pre-university education (gymnasium) after year 12. This is generally seen as an educational milestone and a personal goal for most adolescents. It may therefore be that this achievement is primarily driven by factors related to cognitive capacity, whereas further education beyond 12 years is more determined by other factors including choice of career. These education effects did not diminish with increasing age, corresponding to previous findings (Ritchie et al., 2015; Ritchie & Tucker-Drob, 2018). In contrast, BNT, Cube Drawing Test and one Stroop Test time measure demonstrated a more pronounced benefit of education later in life through interactions between age and education. These results indicate that higher educational attainment might counteract age-related decline. A finding that potentially reflects some form of cognitive reserve (Stern, 2002; Stern & Barulli, 2019).

A differential effect of sex was identified in five tests. In the memory domain, women performed better on the verbal memory tests RAVLT and CERAD, while men exhibited enhanced performance in the Corsi Span visual memory test. In the executive domain, women outperformed men in one of the time measures of the Stroop Test and men outperformed women in the Complex RT Test. These findings concurred with the theories that women often display better language skills (Adani & Capanec, 2019) and that men have an

advantage in visual-spatial working memory (Voyer et al., 2017). It should however be noted that these findings could be highly dependent on culture and situation since sex differences seldom remain undisputed in other populations (Miller & Halpern, 2014; Wallentin, 2009).

The “throughput” index scores suggested by Thorne (2006) and applied in this study, taking both accuracy and speed into account, in most cases were found to add to the variance explained as compared to the separate accuracy and speed variables included in the indices. The only exception was Complex RT, where fewer inhibition errors seemed to compensate for slower reaction times. These findings confirm that the throughput measure is a more sensitive index of performance than either speed or accuracy alone under conditions in which both speed and accuracy decline (or improve).

Only the residuals of slightly fewer than half of the test result measures were considered to be normally distributed. Transformation to normalized residuals was not considered as this may bias model estimates (Schmidt & Finan, 2018). In smaller samples, violation of normality can lead to over or underestimation of the sample variance (Crawford et al., 2006). However, as pointed out by Testa et al. (2009), violation of normality is not a serious violation in the large samples found in this study, and nonnormal residuals do not lead to biased regression coefficients. A more serious violation of the assumptions for regression-based norms is the presence of heteroskedasticity (i.e., when the variance of residuals differs over the continuum of test scores). To account for the presence of heteroskedasticity, we followed the suggested method to stratify the sample by level of predicted performance and compute the standard deviation of the residual for each stratification group (Van der Elst et al., 2006). Even though this method creates somewhat arbitrary groups for different levels of variance, it is still considered the optimal solution to deal with heteroskedasticity (Testa et al., 2009).

Some limitations of this study should be mentioned. First, the currently presented norms are not adjusted for potential

Table 4. SD of residuals for all test results

Cognitive test	Predicted score	SD	Cognitive test	Predicted score	SD
TMT part A index	≤.612	.174	Clock Drawing Test	Total range	.280
	.613–.795	.198	TMT part B index	≤.143	.058
	.796–.982	.246		.144–.260	.106
	≥.983	.318		.261–.375	.132
Time to completion (s)	≤28.921	12.083		≥.376	.162
	28.922–37.777	11.776	Time to completion (s)	≤67.141	31.329
	37.777–46.496	13.282		67.142–90.627	31.534
	≥46.497	21.511		90.628–113.148	34.674
Correct connections	Total range	.613		≥113.149	60.855
SDMT	Total range	6.888	Correct connections	≤19.042	8.788
Simple RT Test index	Total range	2.950		19.043–20.860	5.234
Reaction time (ms)	Total range	84.314		20.860–22.563	3.788
Correct responses	Total range	.391		≥22.564	3.336
RAVLT learning	Total range	7.765	Complex RT Test index	Total range	2.947
RAVLT short-term recall	Total range	2.701	Reaction time (ms)	Total range	118.351
RAVLT long-term recall	Total range	2.817	Correct responses	Total range	.530
RAVLT recognition index	≤11.457	3.450	Inhibition errors	≤.215	.444
	11.458–12.553	2.561		.216–.323	.473
	12.554–13.648	1.861		.323–.444	.703
	≥13.649	2.166		≥.445	.734
Hits	≤13.392	1.788	PASAT	≤42.417	14.987
	13.393–13.874	1.446		42.418–46.174	11.296
	13.875–14.365	1.365		46.175–49.744	9.354
	≥14.366	.819		≥49.745	7.499
False alarms	≤1.274	1.561	Stroop Test index	Total range	2.770
	1.275–2.023	1.973	Inhibition (ms)	≤344.755	272.933
	2.023–2.694	2.460		344.756–418.330	415.199
	≥2.695	2.875		418.330–491.904	351.719
CERAD learning	Total range	3.243		≥491.905	607.320
CERAD recall	Total range	1.894	Reaction time W (ms)	≤1274.136	200.325
CERAD recognition index	Total range	.984		1274.137–1585.312	337.624
Hits	Total range	.722		1585.313–1907.343	591.386
False alarms	Total range	.489		≥1907.344	819.021
Corsi Span forward	Total range	1.010	Correct responses W	Total range	.444
Span backward	Total range	1.085	Reaction time CW (ms)	≤1708.277	381.411
Token Test	Total range	3.006		1708.278–2138.687	397.850
BNT	≤9.576	2.833		2138.688–2547.231	729.708
	9.577–10.083	2.775		≥2547.232	769.834
	10.084–12.575	2.564	Correct responses CW	Total range	1.451
	≥12.576	1.926	Tower of Hanoi index	≤.578	.337
FAS Test	Total range	13.577		.579–.736	.386
Cube Drawing Test	≤.506	.552		.736–.895	.417
	.507–.695	.513		≥.896	.527
	.696–.877	.335	Correct solution	Total range	.137
	≥.878	.256	Time to completion (min)	Total range	1.760

The predicted score is the test takers estimated normative score and can be calculated with Table 3. Total range = any predicted score (SD is not dependent on the predicted score for this test result); TMT = Trail Making Test; SDMT = Symbol Digit Modalities Test; RT = reaction time; RAVLT = Rey Auditory Verbal Learning Test; CERAD = Word List Learning Test; BNT = Boston Naming Test; FAS = Word Fluency Test; PASAT = Paced Auditory Serial Addition Test; Stroop Test W = congruent word trial; Stroop Test CW = incongruent colourword trial.

learning effects which could occur with repeated testing. To do that, a test–retest study is warranted. However, to reduce learning effects, alternate lists are available for RAVLT, BNT and SDMT (alternate lists were not used in this study). Second, the regression coefficients and SD of residuals were calculated treating the normative sample statistics as population statistics. We suggested calculation of

standardized values according to z-distributions. An alternative method would be to employ a t-distribution, as detailed in the Supplementary Methods. However, with sample sizes exceeding 100, as seen in this study, the outcomes of both methods are highly comparable (Crawford & Garthwaite, 2006). Third, the RAVLT and BNT results are based directly on automated speech recognition, without manual result

validation which was not available at the time. In the CERAD and the FAS Test, also dependent on speech, it was observed that a subset of participants was partly misunderstood by the automated speech recognition. Therefore, these individuals received a higher score after manual result validation. If this phenomenon also occurred for RAVLT and BNT, the norms for these tests could be slightly low in our sample compared to the true population. For RAVLT this was tested in a separate unpublished data set: an average of 96% of words were correctly recognised by the automated speech recognition procedure. Nevertheless, the norms for these tests should be used with certain caution. A fourth limitation is that the sample in this study is not selected on optimal health as common age-related diseases were not part of the exclusion criteria. Including non-optimally healthy elderly in the sample implies a risk for underestimating truly healthy functioning in old age (Bergman & Almkvist, 2015). On the other hand, this is common in normative studies (e.g. the studies from the Mayo clinic; Ivnik et al., 1990, 1992, 1996). It should further be mentioned that the present norms were derived from a population that, with few exceptions, were ethnically and culturally Swedish adults who live in economically stable urban regions in Sweden. Therefore, we cannot readily conclude that the findings of the present study will generalise to populations composed of other ethnic, cultural, social or economic backgrounds. There are, however, indications that at least some of the included tests are culture independent (Cores et al., 2015; Møller et al., 1998; Shulman, 2000; Van der Elst et al., 2005). While keeping these limitations in mind, the findings of this study can be applied for the Swedish Mindmore tests and potentially also in other countries.

Some strengths of the present norms are that they are based on a large data set, cover all adult ages, education and sex which are the most influential determinants of cognitive impairment (Lezak et al., 2012). The wide range of included tests covers the key domains attention and processing speed, memory, language, visuospatial functions and executive functions. Therefore, the screening battery is able to assess cognitive impairment associated with a multitude of mental disorders, including common disorders such as MCI and incipient dementia (Levey et al., 2006; Winblad et al., 2004), depression (Hammar & Årdal, 2009; Zaremba et al., 2019) and burnout (Ellbin et al., 2018; Grossi et al., 2015). Finally, the presented normative models may be of more general interest to clinical neuropsychologists as computer-based cognitive assessment gains ground in primary and specialist care (Miller & Barr, 2017; Sabbagh, Boada, Borson, Doraiswamy, et al., 2020).

Comparing patients' results to the norms will guide clinicians using the Mindmore tests in the clinical interpretation of their patients' results. Comparison to normative data provides a measure of standardised deviation of a patient's observed result to their expected result had they been part of the norm population. A common method is to consider negative deviation in performance as less-than-typical when below -1 SD, as suspected impairment when below -1.5 SD,

and as atypical/pathological when below -2 SD (American Psychiatric Association, 2013). To determine the clinical validity of the tests included in the present study, we shall follow-up with separate studies on sensitivity in different clinical populations.

Under certain conditions, such as during an epidemic, it might be preferable to administer a cognitive screening remotely instead of in-person, even with the obvious disadvantages of reduced patient contact and control over the test setting. Other cognitive test batteries, such as the NIH Toolbox-Cognition, have adapted their test protocols for remote administration (Rebchuk et al., 2019). Mindmore is fully self-administrative and as such requires little adaptation to be performed in a remote setting (Owens et al., 2020). Future studies may support the validity of Mindmore in a remote setting.

In conclusion, we collected normative data for 14 traditional cognitive tests, digitised and adapted for self-administration through the Mindmore digital platform. With this normative database, the usability of the Mindmore screening battery is further improved as it will allow for more accurate interpretation of test results, hopefully leading to improved clinical decision making and better care for patients with cognitive impairment.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit <https://doi.org/10.1017/S135561772100045X>

ACKNOWLEDGEMENTS

The authors thank the participants involved in the present study who made this work possible. The authors also acknowledge the great work of the research assistants Lina Lindrågen, Adrian Ringström and Ludwig Franke Föyen, the study staff and Jan J. Kuiper.

FINANCIAL SUPPORT

This work was supported by Vinnova (Innovative startups step 2, grant number 2018-02179).

CONFLICTS OF INTEREST

Wobbe van den Hurk, Ingvar Bergman and Alejandra Machado are or have been employees of Mindmore AB. Anders Gustavsson indirectly owns shares in Mindmore AB. Jonas Bjeremo declared no conflicts of interest.

REFERENCES

- Aarsland, D., Creese, B., Politis, M., Chaudhuri, K. R., Ffytche, D. H., Weintraub, D., & Ballard, C. (2017). Cognitive decline in Parkinson disease. *Nature Reviews Neurology*, *13*(4), 217–231. <https://doi.org/10.1038/nrneuro.2017.27>

- Adani, S., & Cepanec, M. (2019). Sex differences in early communication development: Behavioral and neurobiological indicators of more vulnerable communication system development in boys. *Croatian Medical Journal*, *60*(2), 141–149. <https://doi.org/10.3325/cmj.2019.60.141>
- American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). American Psychiatric Publishing Inc.
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). American Psychiatric Publishing Inc.
- Andrés, P., Parmentier, F. B. R., & Escera, C. (2006). The effect of age on involuntary capture of attention by irrelevant sounds: A test of the frontal hypothesis of aging. *Neuropsychologia*, *44*(12), 2564–2568. <https://doi.org/10.1016/j.neuropsychologia.2006.05.005>
- Army Individual Test Battery. (1944). *Manual of directions and scoring*. Washington, DC: War Department, Adjutant General's Office.
- Bauer, R. M., Iverson, G. L., Cernich, A. N., Binder, L. M., Ruff, R. M., & Naugle, R. I. (2012). Computerized neuropsychological assessment devices: Joint position paper of the American academy of clinical neuropsychology and the national academy of neuropsychology. *Archives of Clinical Neuropsychology*, *27*(3), 362–373. <https://doi.org/10.1093/arclin/acs027>
- Benton, A. L., Hamsher, K. D., & Sivan, A. (1989). *Multilingual Aphasia Examination*. Iowa City, IA: AJA Associates. *NEUROPSYCHOLOGY, BLOCKING, SCHIZOPHRENIA*, *59*.
- Bergman, I., & Almkvist, O. (2015). Neuropsychological test norms controlled for physical health: Does it matter? *Scandinavian Journal of Psychology*, *56*(2), 140–150. <https://doi.org/10.1111/sjop.12170>
- Bergman, I., Johansson, K., Almkvist, O., & Lundberg, C. (2016). Health-adjusted neuropsychological test norms based on 463 older Swedish car drivers. *Scandinavian Journal of Psychology*, *57*(2), 93–107. <https://doi.org/10.1111/sjop.12273>
- Binder, L. M., Iverson, G. L., & Brooks, B. L. (2009). To err is human: “abnormal” Neuropsychological scores and variability are common in healthy adults. *Archives of Clinical Neuropsychology*, *24*(1), 31–46. <https://doi.org/10.1093/arclin/acn001>
- Björngrim, S., Van den Hurk, W., Betancort, M., Machado, A., & Lindau, M. (2019). Comparing traditional and digitized cognitive tests used in standard clinical evaluation – a study of the digital application Minnamera. *Frontiers in Psychology*, *10*(2327), 1–11. <https://doi.org/10.3389/fpsyg.2019.02327>
- Boller, F., & Vignolo, L. A. (1966). Latent sensory aphasia in hemisphere-damaged patients: an experimental study with the Token Test. *Brain*, *89*(4), 815–830. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=5335067
- Carter, S. F., Caine, D., Burns, A., Herholz, K., & Ralph, M. A. L. (2012). Staging of the cognitive decline in Alzheimer's disease: Insights from a detailed neuropsychological investigation of mild cognitive impairment and mild Alzheimer's disease. *International Journal of Geriatric Psychiatry*, *27*(4), 423–432. <https://doi.org/10.1002/gps.2738>
- Chowdhury, M. Z. I., & Turin, T. C. (2020). Variable selection strategies and its importance in clinical prediction modelling. *Family Medicine and Community Health*, *8*, e000262. <https://doi.org/10.1136/fmch-2019-000262>
- Cohen, R. A., Marsiske, M. M., & Smith, G. E. (2019). Neuropsychology of aging. In S. T. DeKosky & S. Asthana (Eds.), *Handbook of Clinical Neurology* (3rd series, Vol. 167, pp. 149–180). Elsevier. <https://doi.org/10.1016/B978-0-12-804766-8.00010-8>
- Cores, E. V., Vanotti, S., Eizaguirre, B., Fiorentini, L., Garcea, O., Benedict, R. H. B., & Cáceres, F. (2015). The effect of culture on two information-processing speed tests. *Applied Neuropsychology: Adult*, *22*(4), 241–245. <https://doi.org/10.1080/23279095.2014.910214>
- Corsi, P. M. (1972). Human memory and the medial temporal region of the brain. In *Department of Psychology*. McGill University.
- Crawford, J. R., & Garthwaite, P. H. (2006). Comparing patients' predicted test scores from a regression equation with their obtained scores: A significance test and point estimate of abnormality with accompanying confidence limits. *Neuropsychology*, *20*(3), 259–271. <https://doi.org/10.1037/0894-4105.20.3.259>
- Crawford, J. R., Garthwaite, P. H., Azzalini, A., Howell, D. C., & Laws, K. R. (2006). Testing for a deficit in single-case studies: Effects of departures from normality. *Neuropsychologia*, *44*(4), 666–677. <https://doi.org/10.1016/j.neuropsychologia.2005.06.001>
- Crawford, J. R., & Howell, D. C. (1998). Regression equations in clinical neuropsychology: An evaluation of statistical methods for comparing predicted and obtained scores. *Journal of Clinical and Experimental Neuropsychology*, *20*(5), 755–762. <https://doi.org/10.1076/jcen.20.5.755.1132>
- Critchley, M. (1953). *The parietal lobes*. Hafner, New York.
- De Renzi, E., & Vignolo, L. A. A. (1962). The token test: A sensitive test to detect receptive disturbances in aphasics. *Brain*, *85*(4), 665–678. <https://doi.org/10.1093/brain/85.4.665>
- Ellbin, S., Engen, N., Jonsdottir, I. H., & Nordlund, A. I. K. K. (2018). Assessment of cognitive function in patients with stress-related exhaustion using the Cognitive Assessment Battery (CAB). *Journal of Clinical and Experimental Neuropsychology*, *40*(6), 567–575. <https://doi.org/10.1080/13803395.2017.1388359>
- Feenstra, H. E., Vermeulen, I. E., Murre, J. M., & Schagen, S. B. (2018). Online Self-Administered cognitive testing using the Amsterdam cognition scan: establishing psychometric properties and normative data. *Journal of Medical Internet Research*, *20*(5), e192. <https://doi.org/10.2196/jmir.9298>
- Finkel, D., Reynolds, C. A., McArdle, J. J., Gatz, M., & Pedersen, N. L. (2003). Latent growth curve analyses of accelerating decline in cognitive abilities in late adulthood. *Developmental Psychology*, *39*(3), 535–550. <https://doi.org/10.1037/0012-1649.39.3.535>
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198.
- Gautam, P., Cherbuin, N., Sachdev, P. S., Wen, W., & Anstey, K. J. (2011). Relationships between cognitive function and frontal grey matter volumes and thickness in middle aged and early old-aged adults: The PATH Through Life Study. *NeuroImage*, *55*(3), 845–855. <https://doi.org/10.1016/j.neuroimage.2011.01.015>
- Germine, L., Reinecke, K., & Chaytor, N. S. (2019). Digital neuropsychology: Challenges and opportunities at the intersection of science and software. *Clinical Neuropsychologist*, *33*(2), 271–286. <https://doi.org/10.1080/13854046.2018.1535662>

- Google Cloud Speech-to-Text. (2020). <https://cloud.google.com/speech-to-text/>
- Gronwall, D. M. A. (1977). Paced auditory serial-addition task: a measure of recovery from concussion. *Percept Mot Skills*, 44(2), 367–373. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=866038
- Gronwall, D. M. A., & Sampson, H. (1974). *The psychological effects of concussion*. Auckland: Oxford University Press.
- Grossi, G., Perski, A., Osika, W., & Savic, I. (2015). Stress-related exhaustion disorder - clinical manifestation of burnout? A review of assessment methods, sleep impairments, cognitive disturbances, and neuro-biological and physiological changes in clinical burnout. *Scandinavian Journal of Psychology*, 56(6), 626–636. <https://doi.org/10.1111/sjop.12251>
- Grzegorski, T., & Losy, J. (2017). Cognitive impairment in multiple sclerosis - A review of current knowledge and recent research. *Reviews in the Neurosciences*, 28(8), 845–860. Walter de Gruyter GmbH. <https://doi.org/10.1515/revneuro-2017-0011>
- Gustavson, D. E., Elman, J. A., Sanderson-Cimino, M., Franz, C. E., Panizzon, M. S., Jak, A. J., . . . Kremen, W. S. (2020). Extensive memory testing improves prediction of progression to MCI in late middle age. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 12(1), 1–9. <https://doi.org/10.1002/dad2.12004>
- Gustavsson, A., Svensson, M., Jacobi, F., Allgulander, C., Alonso, J., Beghi, E., . . . Olesen, J. (2011). Cost of disorders of the brain in Europe 2010. *European Neuropsychopharmacology*, 21(10), 718–779. <https://doi.org/10.1016/j.euroneuro.2011.08.008>
- Hammar, Å., & Årdal, G. (2009). Cognitive functioning in major depression – a summary. *Frontiers in Human Neuroscience*, 3, 26. <https://doi.org/10.3389/neuro.09.026.2009>
- Hlávka, J. P., Mattke, S., & Liu, J. L. (2018). Assessing the Preparedness of the Health Care System Infrastructure in Six European Countries for an Alzheimer's Treatment. In *RAND corporation*. <https://doi.org/10.1017/CBO9780511793677.003>
- Iracleous, P., Nie, J. X., Tracy, C. S., Moineddin, R., Ismail, Z., Shulman, K. I., & Upshur, R. E. G. (2010). Primary care physicians' attitudes towards cognitive screening: Findings from a national postal survey. *International Journal of Geriatric Psychiatry*, 25(1), 23–29. <https://doi.org/10.1002/gps.2293>
- Ivnik, R. J., Malec, J. F., Smith, G. E., Tangalos, E. G., & Petersen, R. C. (1996). Neuropsychological tests' norms above age 55: COWAT, BNT, MAE Token, WRAT-R Reading, AMNART, STROOP, TMT, and JLO. *Clinical Neuropsychologist*, 10(3), 262–278. <https://doi.org/10.1080/13854049608406689>
- Ivnik, R. J., Malec, J. F., Smith, G. E., Tangalos, E. G., Petersen, R. C., Kokmen, E., & Kurland, L. T. (1992). Mayo's older Americans normative studies: WAIS-R norms for ages 56 to 97. *Clinical Neuropsychologist*, 6(SUPPL.), 1–30. <https://doi.org/10.1080/13854049208401877>
- Ivnik, R. J., Malec, J. F., Tangalos, E. G., Petersen, R. C., Kokmen, E., & Kurland, L. T. (1990). The Auditory-Verbal Learning Test (AVLT): Norms for ages 55 years and older. *Psychological Assessment*, 2(3), 304–312. <https://doi.org/10.1037/1040-3590.2.3.304>
- James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., . . . Collaborators, P. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 Diseases and Injuries for 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. In *The Lancet* (Vol. 392, Issue 10159). Lancet Publishing Group. [https://doi.org/10.1016/S0140-6736\(18\)32279-7](https://doi.org/10.1016/S0140-6736(18)32279-7)
- Jørgensen, K., Johannsen, P., & Vogel, A. (2017). A Danish adaptation of the Boston naming test: Preliminary norms for older adults and validity in mild Alzheimer's disease. *Clinical Neuropsychologist*, 31(Supp. 1), 72–87. <https://doi.org/10.1080/13854046.2017.1371337>
- Kaplan, E. F., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test* (2nd ed.). Auckland: Auckland University Press.
- Keesara, S., Jonas, A., & Schulman, K. (2020). Covid-19 and health care's digital revolution. *New England Journal of Medicine*, 382(23), e82. <https://doi.org/10.1056/NEJMp2005835>
- Knight, R. G., McMahon, J., Green, T. J., & Skeaff, C. M. (2006). Regression equations for predicting scores of persons over 65 on the Rey Auditory Verbal Learning Test, the mini-mental state examination, the trail making test and semantic fluency measures. *British Journal of Clinical Psychology*, 45(3), 393–402. <https://doi.org/10.1348/014466505X68032>
- Lakens, D. (2017). Equivalence Tests. *Social Psychological and Personality Science*, 8(4), 355–362. <https://doi.org/10.1177/1948550617697177>
- Levey, A., Lah, J., Goldstein, F., Steenland, K., & Bliwise, D. (2006). Mild cognitive impairment: An opportunity to identify patients at high risk for progression to Alzheimer's disease. In *Clinical Therapeutics* (Vol. 28, Issue 7, pp. 991–1001). <https://doi.org/10.1016/j.clinthera.2006.07.006>
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological Assessment* (5th ed.). New York: Oxford University press.
- Lucas, E. [published under the pseudonym: Claus, N.], (1883). *La Tour d'Hanoi: Véritable casse-tête annamite*. Tours, France: P. Bousrez.
- Maruta, C., Guerreiro, M., De Mendonça, A., Hort, J., & Scheltens, P. (2011). The use of neuropsychological tests across Europe: The need for a consensus in the use of assessment tools for dementia. *European Journal of Neurology*, 18(2), 279–285. <https://doi.org/10.1111/j.1468-1331.2010.03134.x>
- McDonald, B. (2002). A Teaching Note on Cook's Distance-A Guideline. In *Research Letters in the Information and Mathematical Sciences* (Vol. 3). <http://www.massey.ac.nz/~wwiims/research/letters/>
- McDowd, J., & Shaw, R. (2000). Attention and aging: a functional perspective. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of Aging and Cognition* (2nd ed., pp. 221–292). Erlbaum. https://scholarworks.merrimack.edu/psy_facpub/13
- Meijer, W. A., De Groot, R. H. M., Van Gerven, P. W. M., Van Bostel, M. P. J., & Jolles, J. (2009). Level of processing and reaction time in young and middle-aged adults and the effect of education. *European Journal of Cognitive Psychology*, 21(2/3), 216–234. <https://doi.org/10.1080/09541440802091780>
- Menkes, M. W., Armstrong, K., Blackford, J. U., Heckers, S., & Woodward, N. D. (2019). Neuropsychological functioning in early and chronic stages of schizophrenia and psychotic bipolar disorder. *Schizophrenia Research*, 206, 413–419. <https://doi.org/10.1016/j.schres.2018.10.009>
- Miller, D. I., & Halpern, D. F. (2014). The new science of cognitive sex differences. *Trends in Cognitive Sciences*, 18(1), 37–45. <https://doi.org/10.1016/j.tics.2013.10.011>
- Miller, J. B., & Barr, W. B. (2017). The Technology Crisis in Neuropsychology. *Archives of Clinical Neuropsychology*, 32, 541–554. <https://doi.org/10.1093/arclin/acx050>
- Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin*, 27(3),

- 272–277. <https://doi.org/https://doi.org/10.1093/oxfordjournals.bmb.a070866>
- Mindmore screening battery. (2020). *Användarmanual [User manual]* (1.1.5).
- Mitrushina, M. N., Boone, K. B., Razani, J., & D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment* (2nd ed.). New York: Oxford University Press.
- Møller, J. T., Cluitmans, P., Rasmussen, L. S., Houx, P., Rasmussen, H., Canet, J., ... Gravenstein, J. S. (1998). Long-term postoperative cognitive dysfunction in the elderly: ISPOCD1 study. *The Lancet*, *351*, 857–861. [https://doi.org/10.1016/S0140-6736\(97\)07382-0](https://doi.org/10.1016/S0140-6736(97)07382-0)
- Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., Mellits, E. D., & Clark, C. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, *39*(9), 1159–1165. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=2771064
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- Necker, L. A. (1832). Observations on some remarkable optical phenomena seen in Switzerland, and on an optical phenomenon which occurs on viewing a figure of a crystal or geometrical solid. *London and Edinburgh Philosophical Magazine and Journal of Science*, *1*(5), 329–337.
- Newman, C. G. J., Bevins, A. D., Zajicek, J. P., Hodges, J. R., Vuilleumoz, E., Dickenson, J. M., ... Noad, R. F. (2018). Improving the quality of cognitive screening assessments: ACE mobile, an iPad-based version of the Addenbrooke's Cognitive Examination-III. *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring*, *10*, 182–187. <https://doi.org/10.1016/j.dadm.2017.12.003>
- Oosterhuis, H. E. M., van der Ark, L. A., & Sijtsma, K. (2016). Sample size requirements for traditional and regression-based norms. *Assessment*, *23*(2), 191–202. <https://doi.org/10.1177/1073191115580638>
- Owens, A. P., Ballard, C., Beigi, M., Kalafatis, C., Brooker, H., Lavelle, G., ... Aarsland, D. (2020). Implementing remote memory clinics to enhance clinical care during and after COVID-19. *Frontiers in Psychiatry*, *11*(September). <https://doi.org/10.3389/fpsy.2020.579934>
- Pfeiffer, E. (1975). A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. *Journal of the American Geriatrics Society*, *23*(10), 433–441. <https://doi.org/10.1111/j.1532-5415.1975.tb00927.x>
- R Core Team. (2019). *R: A Language and Environment for Statistical Computing* (3.6.0 (2019-04-26)– “Planting of a Tree”). R Foundation for Statistical Computing. <https://www.r-project.org/>
- Rebchuk, A. D., Deptuck, H. M., O'Neill, Z. R., Fawcett, D. S., Silverberg, N. D., & Field, T. S. (2019). Validation of a Novel Telehealth Administration Protocol for the NIH Toolbox-Cognition Battery. *Telemedicine and E-Health*, *25*(3), 237–242. <https://doi.org/10.1089/tmj.2018.0023>
- Reh, R. K., Dias, B. G., Nelson, C. A., Kaufer, D., Werker, J. F., Kolb, B., ... Hensch, T. K. (2020). Critical period regulation across multiple timescales. *Proceedings of the National Academy of Sciences*, *117*(38), 23242–23251. <https://doi.org/10.1073/pnas.1820836117>
- Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.
- Ritchie, S. J., Bates, T. C., & Deary, I. J. (2015). Is education associated with improvements in general cognitive ability, or in specific skills? *Developmental Psychology*, *51*(5), 573–582. <https://doi.org/10.1037/a0038981>
- Ritchie, S. J., & Tucker-Drob, E. M. (2018). How Much Does Education Improve Intelligence? A Meta-Analysis. *Psychological Science*, *29*(8), 1358–1369. <https://doi.org/10.1177/0956797618774253>
- Sabbagh, M. N., Boada, M., Borson, S., Chilukuri, M., Doraiswamy, P. M., Dubois, B., ... Hampel, H. (2020). Rationale for Early Diagnosis of Mild Cognitive Impairment (MCI) supported by Emerging Digital Technologies. *Journal of Prevention of Alzheimer's Disease*, *7*(3), 158–164. <https://doi.org/10.14283/jpad.2020.19>
- Sabbagh, M. N., Boada, M., Borson, S., Doraiswamy, P. M., Dubois, B., Ingram, J., ... Hampel, H. (2020). Early Detection of Mild Cognitive Impairment (MCI) in Primary Care. *Journal of Prevention of Alzheimer's Disease*, *7*(3), 165–170. <https://doi.org/10.14283/jpad.2020.21>
- Sachdev, P. S., Blacker, D., Blazer, D. G., Ganguli, M., Jeste, D. V., Paulsen, J. S., & Petersen, R. C. (2014). Classifying neurocognitive disorders: The DSM-5 approach. *Nature Reviews Neurology*, *10*(11), 634–642. <https://doi.org/10.1038/nrneurol.2014.181>
- Schmid, M., & Hammar, Å. (2013). A follow-up study of first episode major depressive disorder. Impairment in inhibition and semantic fluency—potential predictors for relapse? *Frontiers in Psychology*, *4*, 633. <https://doi.org/10.3389/fpsy.2013.00633>
- Schmidt, A. F., & Finan, C. (2018). Linear regression and the normality assumption. *Journal of Clinical Epidemiology*, *98*, 146–151. <https://doi.org/10.1016/j.jclinepi.2017.12.006>
- Schmidt, M. (1996). *Rey Auditory-Verbal Learning Test*. Los Angeles: Western Psychological Services.
- Shulman, K. I. (2000). Clock-drawing: Is it the ideal cognitive screening test? *International Journal of Geriatric Psychiatry*, *15*(6), 548–561. [https://doi.org/10.1002/1099-1166\(200006\)15:6<548:AID-GPS242>3.0.CO;2-U](https://doi.org/10.1002/1099-1166(200006)15:6<548:AID-GPS242>3.0.CO;2-U)
- Shulman, K. I., Shedletsky, R., & Silver, I. L. (1986). The challenge of time: Clock-drawing and cognitive function in the elderly. *International Journal of Geriatric Psychiatry*, *1*(2), 135–140. <https://doi.org/10.1002/gps.930010209>
- Smith, A. (1982). *Symbol Digit Modalities Test (SDMT). Manual (revised)*. Los Angeles: Western Psychological Services.
- Socialstyrelsen. (2017). *Vård vid depression och ångestsyndrom*. <https://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/20743/2017-12-4.pdf>
- Socialstyrelsen. (2018). *Vård och omsorg vid demenssjukdom*. <https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/nationella-riktlinjer/2018-3-1.pdf>
- Soederberg Miller, L. M., & Lachman, M. E. (2000). Cognitive Performance and the Role of Control Beliefs in Midlife. *Ageing, Neuropsychology, and Cognition*, *7*(2), 69–85. [https://doi.org/10.1076/1382-5585\(200006\)7:2;1-U;FT069](https://doi.org/10.1076/1382-5585(200006)7:2;1-U;FT069)
- Solé, B., Jiménez, E., Torrent, C., Reinares, M., Bonnin, C. del M., Torres, I., ... Vieta, E. (2017). Cognitive Impairment in Bipolar Disorder: Treatment and Prevention Strategies. *International Journal of Neuropsychopharmacology*, *20*(8), 670–680. <https://doi.org/10.1093/ijnp/pyx032>

- Statistiska centralbyrån. (2018a). *Befolkning 16-74 år efter nationell bakgrund, utbildningsnivå och år*. Statistikdatabasen, Befolkningens Utbildning. https://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START_UF_UF0506/UtbsUNBef/
- Statistiska centralbyrån. (2018b). *Befolkning efter ålder och kön*. Statistikdatabasen, Befolkningsstatistik. http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START_BE_BE0101_BE0101A/BefolkningR1860/?rxid=b5c8b781-163a-4ed9-a952-25b2bc508fbf
- Stern, Y. (2002). What Is Cognitive Reserve? Theory and Research Application of the Reserve Concept. *Journal of the International Neuropsychological Society*, 8(3), 448–460. [https://doi.org/10.1017.S1355617701020240](https://doi.org/10.1017/S1355617701020240)
- Stern, Y., & Barulli, D. (2019). Cognitive reserve. In S. T. DeKosky & S. Asthanan (Eds.), *Handbook of Clinical Neurology* (3rd series, Vol. 167, pp. 181–190). Elsevier. <https://doi.org/10.1016/B978-0-12-804766-8.00011-X>
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests, administration, norms, and commentary* (3rd ed.). New York: Oxford University Press.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662.
- StudentKaninen.se. (2020). <https://studentkaninen.se/>
- Svedjebrant, M., & Lindmark, B. (2011). Kartläggning av motorik, kognition och fysisk aktivitetsförmåga hos personer som insjuknat i stroke för 1-2 år sedan samt sambandet mellan dessa variabler. *Uppsala Universitet, Magisterutbildning i Sjukgymnastik*.
- Testa, S. M., Winicki, J. M., Pearlson, G. D., Gordon, B., & Schretlen, D. J. (2009). Accounting for estimated IQ in neuropsychological test performance with regression-based techniques. *Journal of the International Neuropsychological Society*, 15, 1012–1022. <https://doi.org/10.1017/S1355617709990713>
- Thorne, D. R. (2006). Throughput: A simple performance index with desirable characteristics. *Behavior Research Methods*, 38(4), 569–573. <https://doi.org/10.3758/BF03193886>
- Tierney, M. C., & Lerner, M. A. (2010). Computerized cognitive assessment in primary care to identify patients with suspected cognitive impairment. *Journal of Alzheimer's Disease*, 20(3), 823–832. <https://doi.org/10.3233/JAD-2010-091672>
- Troyer, E. A., Kohn, J. N., & Hong, S. (2020). Are we facing a crashing wave of neuropsychiatric sequelae of COVID-19? Neuropsychiatric symptoms and potential immunologic mechanisms. In *Brain, Behavior, and Immunity* (Vol. 87). Academic Press Inc. <https://doi.org/10.1016/j.bbi.2020.04.027>
- Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., & Jolles, J. (2005). Rey's verbal learning test: Normative data for 1855 healthy participants aged 24-81 years and the influence of age, sex, education, and mode of presentation. *Journal of the International Neuropsychological Society*, 11(3), 290–302. <https://doi.org/10.1017/S1355617705050344>
- Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., & Jolles, J. (2006). The stroop color-word test: Influence of age, sex, and education; and normative data for a large sample across the adult age range. *Assessment*, 13(1), 62–79. <https://doi.org/10.1177/1073191105283427>
- Voyer, D., Voyer, S. D., & Saint-Aubin, J. (2017). Sex differences in visual-spatial working memory: A meta-analysis. *Psychonomic Bulletin and Review*, 24(2), 307–334. <https://doi.org/10.3758/s13423-016-1085-7>
- Wallentin, M. (2009). Putative sex differences in verbal abilities and language cortex: A critical review. *Brain and Language*, 108(3), 175–183. <https://doi.org/10.1016/j.bandl.2008.07.001>
- WHO. (2018). *A Vision for Primary Health Care in the 21st Century*. World Health Organisation. http://www.who.int/docs/default-source/primary-health/vision.pdf?sfvrsn=c3119034_2
- WHO. (2019). *Dementia*. World Health Organisation. <https://www.who.int/news-room/fact-sheets/detail/dementia>
- Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L.-O., ... Petersen, R. C. (2004). Mild cognitive impairment - beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *Journal of Internal Medicine*, 256(3), 240–246. <https://doi.org/10.1111/j.1365-2796.2004.01380.x>
- Wingfield, A. (2000). Speech perception and the comprehension of spoken language in adult aging. In D. Park & N. Schwarz (Eds.), *Cognitive Aging: A Primer* (pp. 175–195). Philadelphia: Psychology Press-Taylor & Francis.
- Withall, A., Harris, L. M., & Cumming, S. R. (2009). The relationship between cognitive function and clinical and functional outcomes in major depressive disorder. *Psychological Medicine*, 39(3), 393. <https://doi.org/10.1017/S0033291708003620>
- Wittchen, H. U., Jacobi, F., Rehm, J., Gustavsson, A., Svensson, M., Jönsson, B., ... Steinhausen, H.-C. (2011). The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology*, 21(9), 655–679. <https://doi.org/10.1016/j.euroneuro.2011.07.018>
- Woolley, J. D., Khan, B. K., Murthy, N. K., Miller, B. L., & Rankin, K. P. (2011). The diagnostic challenge of psychiatric symptoms in neurodegenerative disease: Rates of and risk factors for prior psychiatric diagnosis in patients with early neurodegenerative disease. *Journal of Clinical Psychiatry*, 72(2), 126–133. <https://doi.org/10.4088/JCP.10m06382oli>
- Yehuda, R., Tischler, L., Golier, J. A., Grossman, R., Brand, S. R., Kaufman, S., & Harvey, P. D. (2006). Longitudinal Assessment of Cognitive Performance in Holocaust Survivors with and without PTSD. *Biological Psychiatry*, 60(7), 714–721. <https://doi.org/10.1016/j.biopsych.2006.03.069>
- Yoo, W., Mayberry, R., Bae, S., Singh, K., Peter He, Q., & Lillard, J. W. (2014). A Study of Effects of Multi Collinearity in the Multivariable Analysis. *International Journal of Applied Science and Technology*, 4(5), 9–19. <http://www.ncbi.nlm.nih.gov/pubmed/25664257>
- Zahodne, L. B., Stern, Y., & Manly, J. J. (2015). Differing effects of education on cognitive decline in diverse elders with low versus high educational attainment. *Neuropsychology*, 29(4), 649–657. <https://doi.org/10.1037/neu0000141>
- Zaremba, D., Schulze Kalthoff, I., Förster, K., Redlich, R., Grotegerd, D., Leehr, E. J., ... Dannlowski, U. (2019). The effects of processing speed on memory impairment in patients with major depressive disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 92, 494–500. <https://doi.org/10.1016/j.pnpbp.2019.02.015>
- Zimprich, D., & Mascherek, A. (2010). Five views of a secret: Does cognition change during middle adulthood? *European Journal of Ageing*, 7(3), 135–146. <https://doi.org/10.1007/s10433-010-0161-5>
- Zucchella, C., Federico, A., Martini, A., Tinazzi, M., Bartolo, M., & Tamburin, S. (2018). Neuropsychological testing. *Practical Neurology*, 18(3), 227–237. <https://doi.org/10.1136/practneurol-2017-001743>
- Zuckerman, H., Pan, Z., Park, C., Brietzke, E., Musial, N., Shariq, A. S., ... McIntyre, R. S. (2018). Recognition and Treatment of Cognitive Dysfunction in Major Depressive Disorder. *Frontiers in Psychiatry*, 9(655). <https://doi.org/10.3389/fpsy.2018.00655>