


## ORIGINAL ARTICLE

# Cognitive functioning in the general population: Factor structure and association with mental disorders—The neuropsychological test battery of the mental health module of the German Health Interview and Examination Survey for Adults (DEGS1-MH)

Michael Wagner<sup>1,8</sup>  | Steffen Wolfsgrubner<sup>1,8</sup> | Beate Gaertner<sup>2</sup> | Luca Kleineidam<sup>1,8</sup> | Amanda K. Buttery<sup>2,3</sup> | Frank Jacobi<sup>4,5</sup> | Wim Van der Elst<sup>6</sup> | Jelle Jolles<sup>7</sup> | Ulfert Hapke<sup>2</sup> | Hans-Ulrich Wittchen<sup>4</sup> | Wolfgang Maier<sup>1</sup> | Markus A. Busch<sup>2</sup>

<sup>1</sup>Department of Psychiatry and Psychotherapy, University of Bonn, Bonn, Germany

<sup>2</sup>Department of Epidemiology and Health Monitoring, Robert Koch Institute, Berlin, Germany

<sup>3</sup>Faculty of Life Sciences and Medicine, King's College London, London, UK

<sup>4</sup>Institute of Clinical Psychology and Psychotherapy and Center of Clinical Epidemiology and Longitudinal Studies (CELOS), Technische Universität Dresden, Dresden, Germany

<sup>5</sup>Psychologische Hochschule Berlin, Berlin, Germany

<sup>6</sup>I-BioStat, Hasselt, University of Hasselt, Hasselt, Belgium

<sup>7</sup>Department of Educational Neuroscience, Faculty of Psychology and Education and LEARN! Institute, VU University Amsterdam, Amsterdam, The Netherlands

<sup>8</sup>DZNE, German Center for Neurodegenerative Diseases, Bonn, Germany

## Correspondence

Michael Wagner, Department of Psychiatry and Psychotherapy, University of Bonn, Bonn, Germany.

Email: michael.wagner@uni-bonn.de

## Funding information

Stiftung Seelische Gesundheit (DGPPN); Technische Universität Dresden (TUD); German Ministry of Health (Bundesministerium für Gesundheit [BMG])

## Abstract

The objective of this study is to obtain population level data about cognitive functions and their association with mental disorders. We here report factor analytic and psychometric findings of a neuropsychological test battery and examine the association of current and past mental disorders with cognitive function in a large nationwide population-based sample of 18- to 79-year-old adults in Germany ( $n = 3,667$ ) participating in the mental health module of the German Health Interview and Examination Survey for Adults 2008–2011. Confirmatory factor analysis confirmed verbal memory and executive function factors. Older age was strongly associated with lower verbal memory and executive function and with higher vocabulary scores. After adjustment for age, sex, and education, rather modest decrements were found for verbal memory ( $\beta = -.118$ ,  $p = .002$ ) and executive functions ( $\beta = -.191$ ,  $p < .001$ ) in participants with any current mental disorder ( $n = 442$ ) compared to those without ( $n = 3,201$ ). Small decrements in memory ( $\beta = -.064$ ,  $p = .031$ ) and executive function ( $\beta = -.111$ ,  $p < .001$ ) were found in participants with any mental disorder in the last 12 months but not in those with past (fully or partially remitted) mental disorders, compared to participants without a history of mental disorder. More fine-grained analyses of these data will investigate the complex interplay between cognition, health behaviors, and specific mental and somatic diseases.

## KEYWORDS

cognition, epidemiology, neuropsychology, psychometrics

## 1 | INTRODUCTION

Cognitive abilities and their dysfunctions and impairments are increasingly considered important aspects of public health. Cognitive impairments are not only diagnostic criteria for many mental disorders, such as depressive disorders or neurocognitive disorders as defined by Diagnostic and Statistical Manual of Mental Disorders (DSM)-5 (APA, 2013), but they can also be indicators and concomitants of a wide range of neurological and other somatic diseases. They are also discussed as risk factors and markers for the development of many clinical conditions. The assessment of cognitive functions has therefore been recognized as an important element not only for improved clinical management but also for epidemiological research and population health surveys (Deary & Batty, 2007). With regard to mental health, cognitive dysfunctions may explain why mental diseases often are so disabling in terms of functioning in daily life, work productivity, and social interaction (McIntyre & Cha, 2016), having become the most disabling group of diseases in terms of disability-adjusted life years and years lived with disability (Wittchen, Jacobi, Rehm, et al., 2011; Wykes et al., 2015).

However, population-based studies assessing mental health and cognitive function with established psychometric instruments at the same time are rare (Castaneda et al., 2008; Castaneda et al., 2011; Cullen et al., 2015; Simons et al., 2009; Tuulio-Henriksson et al., 2011) and have focused on specific disorders or age groups. Studies across the adult age range combining comprehensive mental health and cognitive assessments are still lacking.

The mental health module of the German Health Interview and Examination Survey for Adults 2009–2012 (DEGS1-MH) allows examining the relationship of cognitive functions and impairment with current or past mental disorders. Towards this goal, we included a neuropsychological test battery to assess cognitive function in a nationwide sample of the general adult population aged 18–79 years in Germany (Jacobi et al., 2015; Jacobi et al., 2013). The test battery was compiled from established individual neuropsychological tests administered to all participants.

Examining the association between individual neuropsychological tests and mental health is a valid and straightforward approach to uncover differential effects on specific cognitive abilities. However, the burden of multiple testing and specific measurement characteristics of the individual tests may also reduce the power to detect effects of mental health on domains of cognition and vice versa. In order to overcome these limitations, information from multiple tests assessing the same domain can be combined in one score that offers a more reliable representation of the cognitive function of interest. For example, this approach has been proposed for improved assessment of different aspects of executive functions and examining their associations to psychopathology (Snyder, Miyake, & Hankin, 2015).

In this paper, we (a) describe the neuropsychological test battery employed in DEGS1-MH; (b) derive and validate cognitive domain scores, report their psychometric properties, and examine associations between cognitive and socio-demographic variables including age, sex, and education; and (c) examine to what degree cognitive functions are associated with the presence of mental disorders.

## 2 | METHODS

### 2.1 | Study design and sample

The design and methods of the German Health Interview and Examination Survey for Adults 2008–2011 (DEGS1) and its DEGS1-MH have been described in detail elsewhere (Gößwald, Lange, Dölle, & Hölling, 2013; Jacobi et al., 2013; Kamtsiuris et al., 2013; Scheidt-Nave et al., 2012). In brief, DEGS1 is a nationwide representative health survey in a population-based sample of  $n = 7,987$  adults aged 18–79 years living in Germany. Persons living in institutions, who were unable to provide written informed consent (e.g., due to severely impaired cognitive function) or who had insufficient German language skills were excluded. A sub-sample of 4,483 adults participated in the DEGS1-MH that collected detailed information on mental disorders and other aspects of mental health (Jacobi et al., 2014; Jacobi et al., 2015; Jacobi et al., 2013), with a median time lag of 6 weeks (interquartile range, 5–25 weeks) between DEGS1 and DEGS1-MH. Data collection in DEGS1-MH included a standardized Composite International Diagnostic Interview (CIDI) for the assessment of mental disorders according to the DSM-IV TR criteria and a neuropsychological test battery (Jacobi et al., 2013). DEGS1 was approved by the federal and state commissioners for data protection and by the ethics committee of Charité-Universitätsmedizin Berlin (no. EA2/047/08). DEGS1-MH was additionally approved by the ethics committee of the Technische Universität Dresden (no. EK174062009). Written informed consent was provided by all participants prior to the interview.

Of 4,483 DEGS1-MH participants, we excluded those who only participated in telephone interviews ( $n = 483$ ), did not complete the neuropsychological test battery ( $n = 94$ ), did not speak German as their first language ( $n = 237$ ), one deaf participant, and one participant with profound learning difficulties. Therefore, the final study sample for this cross-sectional analysis was 3,667 participants (1,894 women and 1,773 men).

### 2.2 | Neuropsychological test battery

The neuropsychological test battery consisted of nine cognitive performance tests (administration time: mean [ $M$ ] = 21.4, standard deviation [ $SD$ ] = 4.3 min) that were selected to assess several domains of cognitive functioning. Test selection was based on psychometric properties (reliability and validity), feasibility, and time burden. The following tests were chosen:

1. *verbal episodic memory* (Trial 1–3 [immediate recall] and Trial 4 [delayed recall] of word lists from the German language version of consortium to establish a registry for Alzheimer's disease [CERAD] measuring the number of correctly recalled words [range 0–10] per trial [Berres, Monsch, Bernasconi, Thalmann, & Stahelin, 2000; Luck et al., 2009; Morris et al., 1989]);
2. *prospective memory* task (remembering two planned actions during the course of the cognitive assessment) from the Cognitive Telephone Screening Instrument (Kliegel, Martin, & Jäger, 2007);
3. *verbal working memory* (Wechsler Intelligence Scale for Adults digit span backwards test measuring the number of digit

sequences of increasing length [2–8 digits] that were correctly recalled in reverse order [range 0–14; Von Aster, Neubauer, & Horn, 2006]);

4. *executive function* and mental speed (verbal fluency task from CERAD measuring the number of animals named within 60 s [Berres et al., 2000; Luck et al., 2009; Morris et al., 1989], letter digit substitution test measuring the number of digits correctly substituted within 60 s [range 0–125; van der Elst, van Boxtel, van Breukelen, & Jolles, 2006], trail making tests [TMT] A and B measuring the time in seconds needed to correctly and accurately connect an ordered sequence of 25 consecutive targets [numbers in TMT-A and alternating numbers and letters in TMT-B; Reitan, 1958]); and a
5. multiple choice vocabulary test (Mehrfachwahl-Wortschatz-Intelligenztest [MWT-B]) measuring the total number of correctly identified real words from 37 short word lists each consisting of one real word and four fictitious pseudowords (Lehrl, 2005). The MWT-B is an established measure of vocabulary and verbal IQ, based on lexical decisions regarding words and similar nonwords.

The number of cognitive dimensions that can be identified in a data set depends on the number and heterogeneity of tests and on sample composition. We aimed to aggregate the test scores into a minimum number of meaningful cognitive domains. A population-based study in the Netherlands (Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2008) with subjects aged 24–81, using a similar set of tests as employed here, found that both a one-factor solution and a two-factor solution of memory versus executive function/speed measures represented the data well. We therefore examined whether such a two-factor solution or even a one-factor solution would be sufficient to describe our data.

## 2.3 | Other variables

Level of education was measured using the Comparative Analysis of Social Mobility in Industrial Nations scheme and classified as low (1a, 1b, and 1c), medium (2a, 2b, and 2c), and high (3a and 3b; Müller, Lüttinger, König, & Karle, 1989), based on self-reported information on school, academic and professional qualifications (Lampert, Kroll, Muters, & Stolzenberg, 2013).

Information on mental disorders were based on the CIDI interview (Jacobi et al., 2014; Jacobi et al., 2013) for DSM-IV TR criteria, covering the following diagnostic groups: anxiety disorders, depressive disorders, bipolar and psychotic disorders, obsessive-compulsive disorders, post-traumatic stress disorders, and substance use disorder. Using the CIDI standard conventions for diagnoses, we defined any mental disorder as meeting the diagnostic criteria for any of the above mentioned diagnoses for three different time frames: current disorder (presence during the last 4 weeks), 12 months diagnosis (presence of disorder during the past 12 months, including 4 weeks), and past diagnoses (diagnostic criteria met not in the past 12 months but at any point in life before). For the respective analyses, 24 participants with a lifetime

history of a mental disorder but with missing data on the recency of any mental disorder were excluded.

## 2.4 | Statistical analysis

Given that we decided to replicate a two-factor solution (see above), we used confirmatory factor analysis (CFA; Kline, 2011) to confirm the a priori assignment of test scores to cognitive domains and to derive aggregated cognitive domain scores. The prospective memory measures were not feasible for CFA due to the limited score range for each of the two prospective tasks (0 point for the failure to initiate any action at the instructed point during the test, 1 point for the initiation of a wrong action at the correct time point, 2 points for the correct prospectively planned action, resulting in a total score ranging from 0 to 4 points). Analyses were carried out in Mplus Version 6.1 (Muthén, 1998–2004) with the full information maximum likelihood (FIML) procedure for parameter estimation and missing data handling (Enders, 2010). We used the robust standard error option to correct for nonnormality. Model fit was evaluated by several absolute and approximate fit statistics including  $\chi^2$  statistic, Akaike's information criterion (AIC), comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean residual (SRMR). CFI values above .95, RMSEA values of .05 or less, and SRMR values below .1 indicate good model fit (Bollen & Long, 1993). Variance and mean of latent factors were fixed to one and zero, respectively.

The following models were fit: (a) a one-factor model that assumes all neuropsychological variables load on a common factor (as we assessed cognitive function on memory and nonmemory domains, we expected insufficient fit for this model); (b) a “two-factor” model with two intercorrelated factors. One factor, called “memory (MEM) factor,” comprised the measures of the verbal episodic memory domain (CERAD word list Trials 1–4). The second factor, called “executive (EXEC) factor,” comprised the variables generally considered to require executive function and mental speed (verbal fluency, letter digit substitution test, TMT-A and TMT-B, and digit span backward). We further investigated intercorrelations of the indicator loading's residuals to check for method correlations, that is, significant correlations among indicators from the same test not captured by the latent construct. Where appropriate, we addressed these by relaxing conditional independence assumptions among the relevant indicators that improves overall model fit (Kline, 2011). Factor score estimates of the latent variables were extracted using the regression method. We computed factor determinacy coefficients to assure that factor scores represent the latent factors adequately. Factor determinacy coefficients >0.90 indicate a reasonable representation whereas scores <0.80 point to an insufficient accordance of factor scores and latent variables (Grice, 2001). Finally, we compared the variance-covariance matrices between persons with a mental disorder during the last 12 months and persons without any mental disorder using  $\chi^2$ -difference testing to test for measurement invariance between groups (Vandenberg & Lance, 2000). A nonsignificant difference would indicate complete invariance in factor structures and latent factor metrics between samples that enables a valid mean comparison.

## 2.5 | Associations of cognitive domain scores and the vocabulary score with factors that affect cognition

We assessed associations of cognitive domain scores and the vocabulary score (MWT-B) with three socio-demographic variables: age, sex, and educational level. We conducted a series of linear regression analyses with nested models in SPSS Version 20. Thirteen participants with missing data on education were excluded from these analyses. In the first model, we included age, sex (reference = male), and educational level (reference = low). In the second model, we additionally added age<sup>2</sup> to assess for quadratic age effects. In the third and fourth model, we additionally examined linear interactions of age and age<sup>2</sup> with sex and educational level, respectively.

In a next step, we evaluated associations between any current mental disorder (past 4 weeks), any 12 months diagnosis (past 12 months, including 4 weeks), and any past diagnosis (anytime during lifetime before 12 months) with cognitive domain and vocabulary scores using linear regression models adjusted for sex, age, age<sup>2</sup>, and educational level. Due to CIDI conventions for diagnoses, current diagnoses and 12-month diagnoses could not be combined in one variable and were analyzed in separate regression models. *p* Values at the 5% level and lower were considered significant.

## 3 | RESULTS

Socio-demographic and neuropsychological characteristics of the study sample are presented in Table 1.

### 3.1 | Confirmatory factor analysis

Results of the CFA showed an insufficient fit for the one-factor model ( $\chi^2 = 3,011$ , degrees of freedom [df] = 27,  $p < .001$ ; AIC = 81,309.810; CFI = .766; RMSEA = .174, 95% confidence intervals [CI]: 0.168–0.179,  $p$  (RMSEA  $\leq .05$ )  $< .001$ ; SRMR = .087). The two-factor solution fitted the data significantly better than the one-factor model and achieved an acceptable model fit ( $\chi^2 = 416$ , df = 26,  $p < .001$ ; AIC = 78,551.390; CFI = .969; RMSEA = .064, 95% CI: 0.059–0.069,  $p$  (RMSEA  $\leq .05$ )  $< .001$ ; SRMR = .034). The MEM and EXEC factors were moderately intercorrelated (see Figure 1). We observed nontrivial correlations for the residuals of the three trials of the CERAD word list immediate recall and for the residuals of TMT-A and TMT-B, respectively. Including these method correlations in the model resulted in considerably improved and excellent fit of the two-factor model to the data ( $\chi^2 = 177$ , df = 26,  $p < .001$ ; AIC = 78,307.908; CFI = .988; RMSEA = .043, 95% CI: 0.037–0.049,  $p$  (RMSEA  $\leq .05$ ) = 0.976; SRMR = .024).

The vocabulary (MWT-B) scores exhibited modest correlations to both factors (Pearson's  $r = .232$ ,  $p < .001$  for MEM;  $r = .268$ ,  $p < .001$  for EXEC), as expected. Prospective memory was more strongly ( $Z = 6.66$ ,  $p > .001$ ) correlated with the EXEC factor (Spearman's  $\rho = .492$ ) than with the MEM factor (Spearman's  $\rho = .414$ ). For cases with complete data, factor determinacy coefficients  $>0.90$  were achieved for MEM ( $\rho = .94$ ) and EXEC ( $\rho = .92$ ) that supports the use of factor score estimates in subsequent analyses (Grice, 2001). However, we observed determinacy coefficients  $<0.80$  for 24 cases (0.7% of the sample) with missing data, resulting in less than three

**TABLE 1** Characteristics of the study sample with neuropsychological assessment ( $n = 3,667$ )

Age (years; M, SD)	52, 7	16, 1
Age group (years; n, %)		
18–24	241	6.57
25–34	319	8.70
35–44	487	13.28
45–54	778	21.22
55–64	731	19.93
65–74	882	24.05
75–81	229	6.24
Female sex (n, %)	1,894	51, 6
Educational level CASMIN ( $n = 3,654$ )		
Low (n, %)	1,122	30, 7
Middle (n, %)	1,799	49, 1
High (n, %)	733	20, 1
Mental disorders ( $n = 3,643$ )		
No mental disorder (n, %)	2,307	63.3
Any mental disorder >12 months (n, %)	413	11.3
Any mental disorder >4 weeks and <12 months (n, %)	481	13.2
Any mental disorder $\leq 4$ weeks (n, %)	442	12.1
Neuropsychological tests		
CERAD WL 1 (M, SD)	5, 83	1, 53
CERAD WL 2 (M, SD)	7, 46	1, 54
CERAD WL 3 (M, SD)	8, 37	1, 42
CERAD WL 4 delayed recall (M, SD)	7, 27	1, 97
CERAD verbal fluency (M, SD)	25, 8	6, 75
LDST (M, SD)	32, 0	7, 97
TMT-A (M, SD)	34, 9	15, 9
TMT-B (M, SD)	84, 3	43, 0
WMS digit span backward (M, SD)	6, 13	1, 96
Prospective memory (M, SD)	3, 02	1, 38
MWT-B (M, SD)	27, 6	4, 92

Note. The table shows descriptive statistics of the study sample of  $N = 3,667$  DEGS-MH participants who completed the face-to-face neuropsychological. CASMIN = Comparative Analysis of Social Mobility in Industrial Nations; CERAD = consortium to establish a registry for Alzheimer's disease; LDST = letter digit substitution test; M = mean; MWT-B = Mehrfachwahl-Wortschatz-Intelligenztest (multiple choice vocabulary test); SD = standard deviation; SES = socioeconomic status; TMT = trail making test; WL (1–4) = word list (Trial 1–4); WMS = Wechsler Memory Scales.

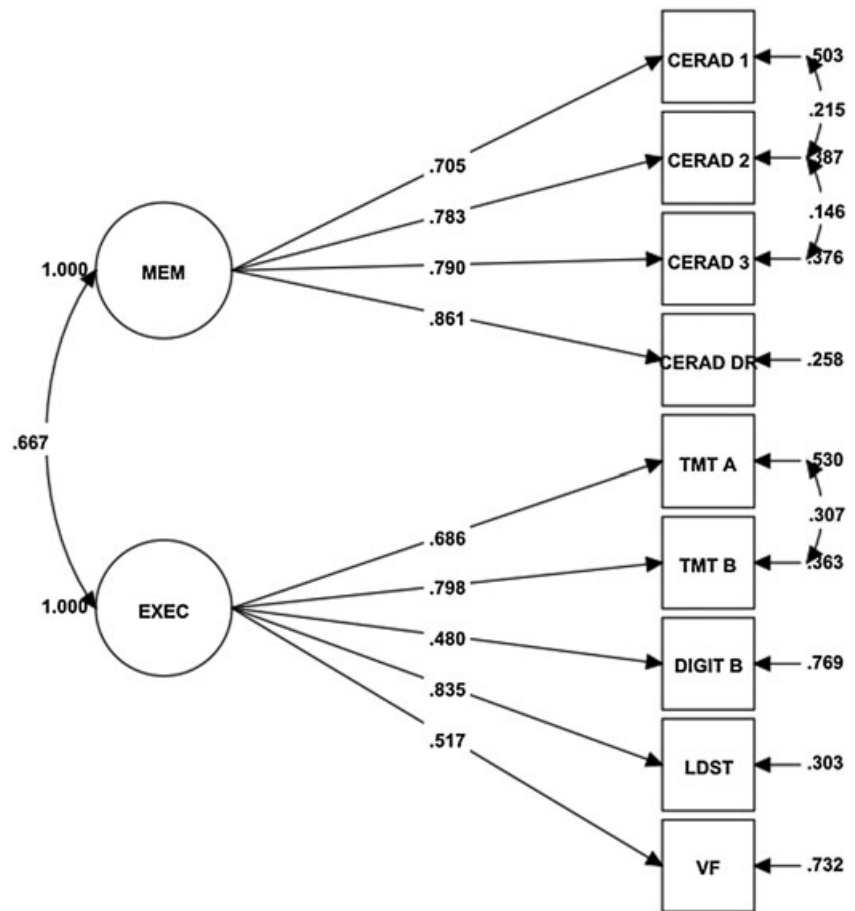
observed values per factor. For these cases, no factor score estimates were computed.

Variance–covariance matrices did not differ between persons with and without any mental disorder in the last 12 months ( $\chi^2 = 27.501$ ; df = 45,  $p = .98$ ) indicating identical factor structures and factor metrics across groups (measurement invariance).

### 3.2 | Associations between socio-demographic variables and cognitive domain scores

Using linear regression analyses (Table 2 and Figure 2a and b), we observed significant effects of age and age<sup>2</sup>, sex (women scored higher

**FIGURE 1** Characteristics of the final two-factor CFA model. One-headed arrows represent factor loadings and residual variances; two headed arrows represent correlations (between factors and residuals, respectively). Values derived from the standardized solution of the final two-factor CFA model. CFA = confirmatory factor analysis; CERAD = consortium to establish a registry for Alzheimer's disease; EXEC = executive function factor; CERAD1–3 = Trials 1–3 of CERAD word list; CERAD4 = delayed recall; DIGIT-B = digit span backward; LDST = letter digit substitution test; MEM = memory factor; MWT-B = Mehrfachwahl-Wortschatz-Intelligenztest (multiple choice vocabulary test); TMT = trail making test; VF = verbal fluency

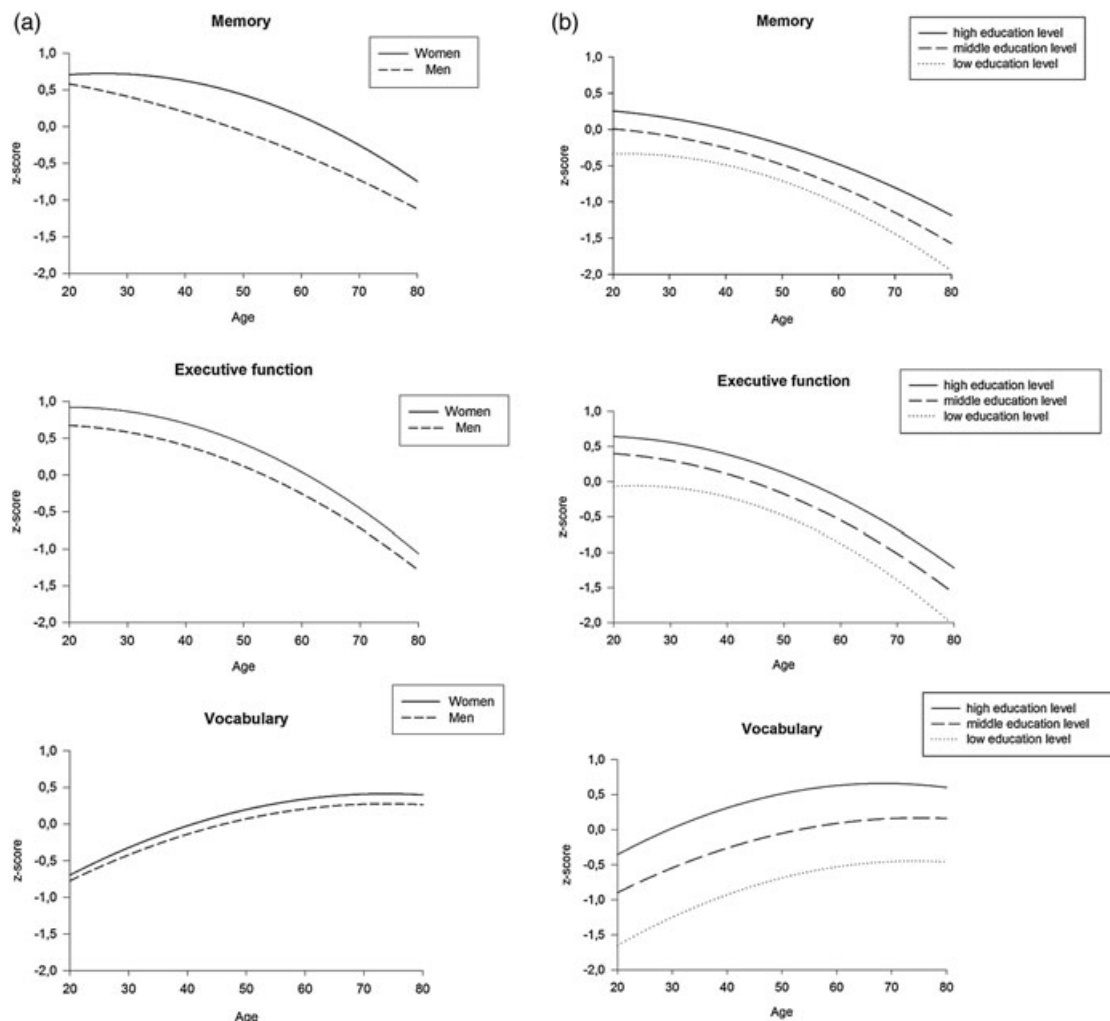


**TABLE 2** Associations of age, sex, and education with cognitive performance in linear regression models

	Memory factor (N = 3,641)		Executive factor (N = 3,628)		Vocabulary score (N = 3,589)	
	$\beta$	<i>p</i>	$\beta$	<i>p</i>	$\beta$	<i>p</i>
<b>Model 1</b>						
Age (beta, <i>p</i> value)	-.439	<.001	-.557	<.001	.300	<.001
Sex (beta, <i>p</i> value)	.238	<.001	.161	<.001	.067	<.001
Middle education (beta, <i>p</i> value)	.160	<.001	.219	<.001	.331	<.001
High education (beta, <i>p</i> value)	.259	<.001	.315	<.001	.487	<.001
Explained variance ( $R^2$ , <i>p</i> value)	.335	<.001	.477	<.001	.226	<.001
<b>Model 2</b>						
Age <sup>2</sup> (beta, <i>p</i> value)	-.633	<.001	-.898	<.001	-.599	<.001
Incremental explained variance ( $R^2$ , <i>p</i> value)	.011	<.001	.022	<.001	.010	<.001
<b>Model 3</b>						
Sex*age (beta, <i>p</i> value)	.973	.001	.234	.355	.092	.770
Sex*age <sup>2</sup> (beta, <i>p</i> value)	-.556	.003	-.163	.327	-.043	.837
Incremental explained variance ( $R^2$ , <i>p</i> value)	.003	<.001	<.001	.600	<.001	.846
<b>Model 4</b>						
Middle education*age (beta, <i>p</i> value)	-.379	.237	-.492	.080	-.197	.574
Middle education*age <sup>2</sup> (beta, <i>p</i> value)	.246	.211	.304	.078	.087	.686
High education*age (beta, <i>p</i> value)	-.434	.297	-.380	.300	-.039	.932
High education*age <sup>2</sup> (beta, <i>p</i> value)	.299	.208	.264	.230	-.031	.911
Incremental explained variance ( $R^2$ , <i>p</i> value)	.001	.374	.001	.302	<.001	.714

*Note.* The table shows results of the linear regression analyses with the two cognitive domain scores (latent factor score estimates derived from the CFA model in Figure 1) and the Mehrfachwahl-Wortschatz-Intelligenztest (multiple choice vocabulary test) raw score as dependent variables and terms of age, age<sup>2</sup>, sex, and educational level as predictors. Beta values are standardized regression coefficients. CFA = confirmatory factor analysis.





**FIGURE 2** (a) Graphical representation of the association of age with cognitive performance in men and women. (b) Graphical representation of the association of age with cognitive performance according to level of education

than men), and education (higher educated participants scored higher than those with lower education) on both cognitive domain scores. For the MWT-B, we observed a significant positive effect of higher education, a *positive* association with age and also a small sex difference (women better than men). We further observed a significant interaction of age and age<sup>2</sup> with sex on the MEM score. The coefficient for the age<sup>2</sup> × sex interaction was positive (i.e., the linear age effect on memory performance is less strong for women) whereas it was negative for the age<sup>2</sup> × sex interaction (i.e., for women the quadratic age effect on memory performance is stronger). This combination of effects is visible in the differently shaped curves in the upper left graph in Figure 2a.

Age, sex, and education together explained 22.6%, 33.6%, and 47.6% of variance in MWT-B, MEM, and EXEC scores, respectively.

### 3.3 | Associations between mental disorders and cognitive domain scores

We examined the associations between any current mental disorder within the past 4 weeks with the two cognitive domains scores (Table 3). After adjustment for age, age<sup>2</sup>, sex, and educational level, significant but rather subtle impairments of MEM ( $\beta = -.118$ ,  $p = .002$ )

and EXEC function ( $\beta = -.191$ ,  $p < .001$ ) were found for current diagnoses compared to no current diagnoses. Associations with cognitive function scores were still significant but with smaller effects estimates for any mental disorder in the last 12 months, whereas a lifetime diagnosis alone (without a diagnosis in the last 12 months) was not at all associated with any impairment (Table 3). The vocabulary score was not associated with current or past diagnoses ( $p > .05$ ) when analyzed likewise.

## 4 | DISCUSSION

We assessed cognitive function in the DEGS1-MH general population sample with a brief battery of established tests and applied CFA to confirm two cognitive domain scores, respectively of verbal memory and executive function. These cognitive domain scores, together with the vocabulary score (based on a single test), are considered to be valuable global indicators of cognition for future analyses with the DEGS1 data set. We first discuss some methodological advantages of our data-reduction approach before addressing the associations with socio-demographic variables and mental disorders and plans for future analyses.

**TABLE 3** Associations of any mental disorder with cognitive performance in linear regression models

	Memory (N = 3,630)		Executive function (N = 3,627)	
	$\beta$	p	$\beta$	p
Current mental disorder (4 weeks)				
Model 1: Adjusted for sex, age, and age <sup>2</sup>				
No current mental disorder (4 weeks)	Ref.		Ref.	
Any current mental disorder (4 weeks)	-.128	.001	-.213	<.001
Model 2: Adjusted for sex, age, age <sup>2</sup> , and educational level				
No current mental disorder (4 weeks)	Ref.		Ref.	
Any current mental disorder (4 weeks)	-.118	.002	-.191	<.001
Mental disorder in the past 12 months				
Model 1: Adjusted for sex, age, and age <sup>2</sup>				
No mental disorder (lifetime)	Ref.		Ref.	
Any mental disorder >12 months	.046	.276	.002	.963
Any mental disorder ≤12 months	-.076	.014	-.129	<.001
Model 2: Adjusted for sex, age, age <sup>2</sup> , and educational level				
No mental disorder (lifetime)	Ref.		Ref.	
Any mental disorder >12 months	.035	.383	-.012	.723
Any mental disorder ≤12 months	-.064	.031	-.111	<.001

#### 4.1 | Comparing the CFA and traditional methods approach

Traditional summary methods to derive composites, such as z-score averaging of subtest scores, are commonly used and are easy to handle. However, these techniques have several shortcomings. In contrast to the z-score averaging method, the CFA approach allows composite scores to be generated using confirmatory testing and single test scores can be meaningfully combined into a composite score. CFA provides measures of overall model fit and information on the fit of individual scores within the model (e.g., factor loadings). Therefore, it is possible to evaluate whether a proposed factor model fits the observed data, and whether computing composite measures are justified. Based on previous evidence (van der Elst et al., 2008), we preferred a two-factor model over a common factor model. Further, our model fit indices demonstrated that the common factor model did not represent the observed cognitive data adequately. In addition, the computation of factor scores ensured that the scores were most highly correlated with the latent variable and offer a reasonable representation of the subject's "true" cognitive abilities. We were also able to identify a small number of cases (0.7% of the sample) for which the computation of factor score estimates was not adequate due to their pattern of missing data. The traditional z-score averaging method does usually not incorporate such a detailed analysis.

#### 4.2 | Associations between socio-demographic variables and cognitive domain scores

Associations found between the three cognitive domains with age, sex, and education are largely in line with prior expectations, indicating lower memory and executive function scores with increasing age. Although cross-sectional data cannot provide a reliable estimate of

longitudinal (individual) cognitive decline because of potentially confounding birth cohort effects (e.g., higher education levels in younger birth cohorts), cross-sectional and longitudinal approaches concur in finding a decrease of most cognitive functions (except vocabulary) already in midlife (Singh-Manoux et al., 2012).

Vocabulary scores were higher in older participants, consistent with other studies (Verhaeghen, 2003), attesting to the experience-dependent "crystalline intelligence" that increases or at least remains stable over the life span. In part, the apparent increase may result from age-cohort influences on vocabulary used in such tests, albeit only a few target items included in the MWT-B are obviously "outdated" (e.g., "Recke," an old German word for warrior). Vocabulary was only modestly correlated with verbal memory and executive function, respectively. As expected from the literature (Dowling, Hermann, La Rue, & Sager, 2010; Van der Elst et al., 2008), verbal memory and executive function were substantially intercorrelated and therefore can be considered as two facets of "fluid intelligence" according to the Cattell-Horn theory of intelligence (Horn, 1989). Prospective memory was more closely related to executive function than to verbal memory, confirming our reasoning that inclusion of prospective memory into the memory domain would be inappropriate.

Our finding of better memory among women is consistent with previous studies describing women's superior performance in tests using auditory verbal stimuli (Pauls, Petermann, & Lepach, 2013). However, associations between gender and cognitive performance may change in the presence of neurodegenerative processes that may affect cognition differentially in men versus women. This has recently been shown for patients with Parkinson's disease (Fengler et al., 2016).

Sex differences were also found in the executive domain with women achieving higher scores than men. A possible explanation for this finding may relate to the inclusion of two timed tasks in this domain that rely on verbal semantic memory (i.e., category fluency)

and verbal working memory (i.e., digit span). A trend for better cognitive performance in elderly women as compared to men, across a range of cognitive tests, has also been noted in another population-based study (Hayat et al., 2014).

Even though educational level was significantly associated with current cognitive performance in our analysis, differences in education explained only a modest proportion (5% variance in memory, 7.5% variance in executive function, and 17.8% variance in vocabulary) of differences in cognitive abilities when added to models that were already adjusted for age and sex. Consequently, educational level should be considered a poor proxy for current cognitive abilities, underscoring the need of measuring cognitive abilities in epidemiological studies.

Consistent with longitudinal studies on cognitive decline, we found that the negative association of memory and executive functions with increasing age is not moderated by education (Singh-Manoux et al., 2011). This may suggest that “cognitive reserve,” which is often inferred from education, cannot mitigate the speed of “normal” age-related cognitive decline until age 80. However, in late life, the average memory and executive performance difference between the highest and lowest educational groups is roughly equivalent to 10 years of age (compare Figure 2b). It is well established that people with higher education experience clinically and functionally significant cognitive impairment (i.e., dementia) much later in their lives (Stern, 2012; Valenzuela & Sachdev, 2006a, 2006b). Consequently, improved education may have a substantial effect on lowering dementia incidence rates in the long term (Norton, Matthews, Barnes, Yaffe, & Brayne, 2014).

### 4.3 | Associations between mental disorders and cognitive domain scores

Despite the fact that cognitive dysfunctions are explicit diagnostic criteria for many mental disorders, the cognitive impairments associated with having any mental disorder were remarkably small (0.1–0.2 SD on average) and were confined to those who met diagnostic criteria within the last 4 weeks. Given substantial evidence from clinical studies that, for example, depressive disorders—that account for almost 40% of our study sample—reveal still persisting cognitive dysfunctions and impairment even months after remission from the episode, this finding was somewhat unexpected, although it is in line with other population-based studies who found no association of lifetime anxiety or depressive disorders with cognition (Castaneda et al., 2008, 2011; Cullen et al., 2015). We assume that this finding is due to three factors. First, our findings are based on a sample of the general household population, likely to underrepresent institutionalized and more severe mental disorder. Thus, subjects who were hospitalized during the field work period because of mental problems or other reasons are not covered. It is also noteworthy that more than 80% of the DEGS1-MH participants with a 12-month diagnosis of mental disorder had not used mental health services because of their condition in the past year (Mack et al. 2014), underlining that population samples of subjects with mental disorders show only limited overlap with samples from clinical settings that usually show more pronounced cognitive impairments. Second, subjects with a mental disorder were compared in this

paper with a random sample of controls without a history of mental disorders. However, we did not screen for the absence of neurological or somatic diseases, as is usually the case in healthy control groups employed in clinical studies. Third, the low level of cognitive functioning decrements observed might be due to the joint effects of (a) using crude categorical measures of diagnostic status, (b) lumping all mental disorders together and not examining differentially the effects of various psychopathological syndromes, and (c) not taking into account their dimensionally measured severity at the time of examination. The finding of a significant, though modest effect, in acute patients with mental disorders might indicate that future detailed analyses might result in more pronounced effects in specific diagnostic subgroups.

The present study, to our best knowledge, is one of the first population health surveys that combine a detailed mental health examination with neuropsychological testing. The detailed analysis of specific (e.g., affective) disorders and syndromes by severity as well as findings for single tests will be presented in follow-up publications. However, the current data suggest that, at the population level, most subjects who suffered from a mental disorder in the past have very limited, if any, enduring cognitive impairments.

### 4.4 | Strengths and limitations

The main strengths of this study are the large nationwide sample of adults aged of 18–79 years from the general population and the assessment of both cognitive functions and mental health status. Measurement invariance testing showed a highly similar factor structure and factor metrics in subjects with and without mental disorders. There are some limitations that should be considered. First, DEGS1 did not include persons living in institutions or those with severe cognitive impairment, and persons with severe medical problems or impaired functional mobility are also likely to be underrepresented. Also, excluded participants without personal neuropsychological assessment and with a non-German first language were on average younger than the included participants, so that the sample of our study was on average older than the total DEGS1-MH sample. However, included and excluded participants did not differ by sex or educational level. Due to the differences in mean age, our study sample might not be representative for the German general population, but our analyses still provide valid factor analytic and psychometric findings of the neuropsychological test battery. Second, our brief battery was not—and was not meant to be—a comprehensive assessment of cognitive functioning, as, for example, reasoning or spatial tasks were not included. This was due to time restrictions and followed from our initial decision to focus on tests with established robust psychometric properties including sensitivity to change for major medical conditions. Third, the 10-item word list of the CERAD was rather easy to learn for younger subjects, resulting in ceiling effects especially for the third learning trial.

### 4.5 | Conclusions and implications for future studies

From the cognitive assessment devised for adults aged 18–79 years in DEGS1-MH, two CFA-derived cognitive factors for verbal memory and



executive function were validated and, together with the vocabulary score, will be used in future analyses of the DEGS1 data set to investigate the complex interplay between cognition, health behaviors, and disease. A substantial proportion of DEGS1 participants with neuropsychological assessment participated in the previous German National Health Interview and Examination Survey 1998 (Bellach, Knopf, & Thefeld, 1998) and its mental health supplement (Jacobi et al., 2004; Jacobi et al., 2002) about 12 years before DEGS1 (Kamtsiuris et al., 2013; Scheidt-Nave et al., 2012), and the cognitive measures described herein will provide new opportunities for longitudinal analyses.

## ACKNOWLEDGEMENTS

DEGS1 was funded primarily by the German Ministry of Health (Bundesministerium für Gesundheit [BMG]). Supplementary funding for DEGS1-MH was provided by the Technische Universität Dresden (TUD) and by the Foundation for Mental Health (Stiftung Seelische Gesundheit) inaugurated by the German Association for Psychiatry, Psychotherapy, and Psychosomatics (Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde [DGPPN]). The neuropsychological section was jointly developed and implemented by TUD (Frank Jacobi), Robert Koch Institute (RKI; Markus Busch), and the University of Bonn (Department of Psychiatry: Michael Wagner).

We would like to acknowledge the support of all RKI staff in the field work of the project, in particular Dr. Heike Hölling for quality assurance, Ingeburg Seiffert for data management, Dr. Stefan Dahm for statistical support, and the responsible RKI project officers for DEGS1-MH, Michael Lange and Dr. Ulfert Hapke. We also thank the many staff members and interviewers of DEGS1-MH at TUD. Above all, we wish to thank all DEGS1-MH study participants.

## DECLARATION OF INTEREST STATEMENT

The authors have no competing interest to declare.

## ORCID

Michael Wagner  <http://orcid.org/0000-0003-2589-6440>

## REFERENCES

- American Psychiatric Association (APA) (2013). *Diagnostic and statistical manual of mental disorders* (Fifth ed.). Arlington, VA: American Psychiatric Association.
- Bellach, B. M., Knopf, H., & Thefeld, W. (1998). The German Health Survey. 1997/98. *Gesundheitswesen (Bundesverband der Ärzte des Öffentlichen Gesundheitsdienstes (Germany))*, 60(Suppl 2), S59–S68.
- Berres, M., Monsch, A. U., Bernasconi, F., Thalmann, B., & Stahelin, H. B. (2000). Normal ranges of neuropsychological tests for the diagnosis of Alzheimer's disease. *Studies in Health Technology and Informatics*, 77, 195–199.
- Bollen, K. A., & Long, J. S. (1993). *Testing structural equation models*. Newbury Park, CA: Sage.
- Castaneda, A. E., Suvisaari, J., Marttunen, M., Perälä, J., Saarni, S. I., Aalto-Setälä, T., ... Tuulio-Henriksson, A. (2008). Cognitive functioning in a population-based sample of young adults with a history of non-psychotic unipolar depressive disorders without psychiatric comorbidity. *Journal of Affective Disorders*, 110(1–2), 36–45.
- Castaneda, A. E., Suvisaari, J., Marttunen, M., Perälä, J., Saarni, S. I., Aalto-Setälä, T., ... Tuulio-Henriksson, A. (2011). Cognitive functioning in a population-based sample of young adults with anxiety disorders. *European Psychiatry*, 26, 346–353.
- Cullen, B., Nicholl, B. I., Mackay, D. F., Martin, D., Ul-Haq, Z., McIntosh, A., ... Smith, D. J. (2015). Cognitive function and lifetime features of depression and bipolar disorder in a large population sample: Cross-sectional study of 143,828 UK Biobank participants. *European Psychiatry*, 30, 950–958. <https://doi.org/10.1016/j.eurpsy.2015.08.006>
- Deary, I. J., & Batty, G. D. (2007). Cognitive epidemiology. *Journal of Epidemiology and Community Health*, 61(5), 378–384.
- Dowling, N. M., Hermann, B., La Rue, A., & Sager, M. A. (2010). Latent structure and factorial invariance of a neuropsychological test battery for the study of preclinical Alzheimer's disease. *Neuropsychology*, 24(6), 742–756.
- Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Publications.
- Fengler, S., Roeske, S., Heber, I., Reetz, K., Schulz, J. B., Riedel, O., ... Hilker, R. (2016). Verbal memory declines more in female patients with Parkinson's disease: The importance of gender-corrected normative data. *Psychological Medicine*, 46(11), 2275–2286.
- Gößwald, A., Lange, M., Döller, R., & Hölling, H. (2013). The first wave of the German Health Interview and Examination Survey for Adults (DEGS1). Recruitment of participants, fieldwork, and quality assurance. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*, 56, 611–619.
- Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods*, 6(4), 430.
- Hayat, S. A., Luben, R., Moore, S., Dalzell, N., Bhaniani, A., Anuj, S., ... Brayne, C. (2014). Cognitive function in a general population of men and women: A cross sectional study in the European Investigation of Cancer–Norfolk cohort (EPIC–Norfolk). *BMC Geriatrics*, 14, 142.
- Horn, J. L. (1989). *Intelligence: Measurement, theory, and public policy* (pp. 29–73). Urbana: Univ. Illinois Press.
- Jacobi, F., Hofler, M., Siegert, J., Mack, S., Gerschler, A., Scholl, L., ... Wittchen, H. U. (2014). Twelve-month prevalence, comorbidity and correlates of mental disorders in Germany: The mental health module of the German Health Interview and Examination Survey for Adults (DEGS1-MH). *International Journal of Methods in Psychiatric Research*, 23(3), 304–319.
- Jacobi, F., Höfler, M., Strehle, J., Mack, S., Gerschler, A., Scholl, L., ... Gaebel, W. (2015). Twelve-months prevalence of mental disorders in the German Health Interview and Examination Survey for Adults–Mental Health Module (DEGS1-MH): a methodological addendum and correction. *International journal of Methods in Psychiatric Research*, 24(4), 305–313.
- Jacobi, F., Mack, S., Gerschler, A., Scholl, L., Hofler, M., Siegert, J., ... Wittchen, H. U. (2013). The design and methods of the mental health module in the German Health Interview and Examination Survey for Adults (DEGS1-MH). *International Journal of Methods in Psychiatric Research*, 22(2), 83–99.
- Jacobi, F., Wittchen, H. U., Holting, C., Hofler, M., Pfister, H., Muller, N., & Lieb, R. (2004). Prevalence, co-morbidity and correlates of mental disorders in the general population: Results from the German Health Interview and Examination Survey (GHS). *Psychological Medicine*, 34(4), 597–611.
- Jacobi, F., Wittchen, H. U., Holting, C., Sommer, S., Lieb, R., Hofler, M., & Pfister, H. (2002). Estimating the prevalence of mental and somatic disorders in the community: Aims and methods of the German National Health Interview and Examination Survey. *International Journal of Methods in Psychiatric Research*, 11(1), 1–18.
- Kamtsiuris, P., Lange, M., Hoffmann, R., Schaffrath Rosario, A., Dahm, S., Kuhnert, R., & Kurth, B. M. (2013). The first wave of the German Health Interview and Examination Survey for Adults (DEGS1): Sample design, response, weighting and representativeness. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*, 56, 620–630.

- Kliegel, M., Martin, M., & Jager, T. (2007). Development and validation of the Cognitive Telephone Screening Instrument (COGTEL) for the assessment of cognitive function across adulthood. *The Journal of Psychology*, 141(2), 147–170.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.
- Lampert, T., Kroll, L., Muters, S., & Stolzenberg, H. (2013). Measurement of socioeconomic status in the German Health Interview and Examination Survey for Adults (DEGS1). *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*, 56(5–6), 631–636.
- Lehrl, S. (2005). *Mehrfachwahl-Wortschatz-Intelligenztest (MWT-B)*. Göttingen: Hogrefe.
- Luck, T., Riedel-Heller, S. G., Wiese, B., Stein, J., Weyerer, S., Werle, J., ... Pentzek, M. (2009). CERAD-NP battery: Age-, gender- and education-specific reference values for selected subtests. Results of the German study on ageing, cognition and dementia in primary care patients (AgeCoDe). *Zeitschrift für Gerontologie und Geriatrie*, 42(5), 372–384.
- Mack, S., Jacobi, F., Gerschler, A., Strehle, J., Höfler, M., Busch, M. A., ... Wittchen, H.-U. (2014). Self-reported utilization of mental health services in the adult German population—evidence for unmet needs? Results of the DEGS1-Mental Health Module (DEGS1-MH). *International Journal of Methods in Psychiatric Research*, 23(3), 289–303.
- McIntyre, R. S., & Cha, D. S. (2016). *Cognitive Impairment in Major Depressive Disorder: Clinical Relevance, Biological Substrates, and Treatment Opportunities*. Cambridge: Cambridge University Press.
- Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., ... 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.
- Müller, W., Lüttinger, P., König, W., & Karle, W. (1989). Class and education in industrial nations. *International Journal of Sociology*, 19, 3–39.
- Muthén, B. O. (1998–2004). *Mplus technical appendices*. Los Angeles, CA: Muthén & Muthén.
- Norton, S., Matthews, F. E., Barnes, D. E., Yaffe, K., & Brayne, C. (2014). Potential for primary prevention of Alzheimer's disease: An analysis of population-based data. *Lancet Neurology*, 13(8), 788–794.
- Pauls, F., Petermann, F., & Lepach, A. C. (2013). Gender differences in episodic memory and visual working memory including the effects of age. *Memory*, 21(7), 857–874.
- Reitan, R. M. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual and Motor Skills*, 8, 271–276.
- Scheidt-Nave, C., Kamtsiuris, P., Goesswald, A., Hoelling, H., Lange, M., Busch, M. A., ... Kurth, B. M. (2012). German Health Interview and Examination Survey for Adults (DEGS)—Design, objectives and implementation of the first data collection wave. *BMC Public Health*, 12(1), 730.
- Simons, C. J. P., Jacobs, N., Derom, C., Thiery, E., Jolles, J., Van Os, J., & Krabbendam, L. (2009). Cognition as predictor of current and follow-up depressive symptoms in the general population. *Acta Psychiatrica Scandinavica*, 120(1), 45–52.
- Singh-Manoux, A., Kivimaki, M., Glymour, M. M., Elbaz, A., Berr, C., Ebmeier, K. P., ... Dugravot, A. (2012). Timing of onset of cognitive decline: Results from Whitehall II prospective cohort study. *BMJ (Clinical research ed)*, 344, d7622.
- Singh-Manoux, A., Marmot, M. G., Glymour, M., Sabia, S., Kivimaki, M., & Dugravot, A. (2011). Does cognitive reserve shape cognitive decline? *Annals of Neurology*, 70(2), 296–304.
- Snyder, H. R., Miyake, A., & Hankin, B. L. (2015). Advancing understanding of executive function impairments and psychopathology: Bridging the gap between clinical and cognitive approaches. *Frontiers in Psychology*, 6.
- Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurology*, 11(11), 1006–1012.
- Tuulio-Henriksson, A., Perälä, J., Saarni, S. I., Isometsä, E., Koskinen, S., Lönnqvist, J., & Suvisaari, J. (2011). Cognitive functioning in severe psychiatric disorders: A general population study. *European Archives of Psychiatry and Clinical Neuroscience*, 261, 447–456.
- Valenzuela, M. J., & Sachdev, P. (2006a). Brain reserve and cognitive decline: A non-parametric systematic review. *Psychological Medicine*, 36(8), 1065–1073.
- Valenzuela, M. J., & Sachdev, P. (2006b). Brain reserve and dementia: A systematic review. *Psychological Medicine*, 36(4), 441–454.
- Van der Elst, W., van Boxtel, M. P., van Breukelen, G. J., & Jolles, J. (2006). The letter digit substitution test: Normative data for 1,858 healthy participants aged 24–81 from the Maastricht Aging Study (MAAS): Influence of age, education, and sex. *Journal of Clinical and Experimental Neuropsychology*, 28(6), 998–1009.
- Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., & Jolles, J. (2008). A large-scale cross-sectional and longitudinal study into the ecological validity of neuropsychological test measures in neurologically intact people. *Archives of Clinical Neuropsychology*, 23(7–8), 787–800.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4–70.
- Verhaeghen, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology and Aging*, 18(2), 332–339.
- Von Aster, M., Neubauer, A., & Horn, R. (2006). *Wechsler-Intelligenztest für Erwachsene. WIE—Manual*. Frankfurt/M: Harcourt Test Services.
- Wittchen, H. U., Jacobi, F., Rehm, J., et al. (2011). The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology*, 21(9), 655–679.
- Wykes, T., Haro, J. M., Belli, S. R., Obradors-Tarrago, C., Arango, C., Ayuso-Mateos, J. L., ... Wittchen, H. U. (2015). Mental health research priorities for Europe. *Lancet Psychiatry*, 2(11), 1036–1042.

**How to cite this article:** Wagner M, Wolfsgruber S, Gaertner B, et al. Cognitive functioning in the general population: Factor structure and association with mental disorders—The neuropsychological test battery of the mental health module of the German Health Interview and Examination Survey for Adults (DEGS1-MH). *Int J Methods Psychiatr Res*. 2018;27:e1594. <https://doi.org/10.1002/mpr.1594>