Supplemental Table 1. Item parameters from unidimensional CFA models: Results from LASI-DAD (N=3,224)

Domain and item	Raw loading	Standardized loading	Thresholds (categorical) or intercept (continuous)							
			1	2	3	4	5	6	7	8
Orientation										
Month	0.77	0.77	-0.93							
Year	0.92	0.92	0.02							
Day of the week	0.58	0.58	-0.96							
Day of the month	0.74	0.74	-0.37							
Season	0.60	0.60	-0.97							
State	0.74	0.74	-0.28							
City	0.66	0.66	-1.59							
Hospital name (or district if at home)	0.73	0.73	-0.75							
Area of town/village or street name	0.68	0.68	-1.16							
Floor of building (or where we are, if at home)	0.43	0.43	-1.31							
Prime Minister	0.78	0.78	-0.22							
Immediate memory										
Three word recall	0.55	0.55	-2.26	-2.17	-2.10	-2.01	-1.95	-1.85	-1.59	-1.40
Ten word recall	0.14	0.73	0.41							
Brave man, immediate recall	0.17	0.69	0.38							
Logical Memory, immediate recall	0.12	0.67	0.18							
Delayed memory										
Ten word delayed recall	0.18	0.77	0.31							
Three word delayed recall	0.55	0.55	-1.04	-0.49	0.25					
Brave man, delayed recall	0.16	0.60	0.21							
Logical Memory, delayed recall	0.16	0.64	0.19							
Constructional Praxis, delayed	0.58	0.58	-0.27	0.55	1.28	2.18				
Recognition memory										
Ten word recognition	0.13	0.55	0.81							
Logical memory, recognition	0.13	0.58	0.53							
Reasoning										
Raven's progressive matrices	0.13	0.63	0.45							

Go-no-go trial 1	0.25	0.72	0.65					
Go-no-go trial 2	0.24	0.68	0.50					
Clock drawing	0.69	0.69	-0.10	0.55	1.12			
Attention/speed								
Serial 7s	0.77	0.77	-0.53	0.08	0.41	0.73	1.08	
Backward day naming	0.78	0.78	-0.72	-0.54	-0.42	-0.33	-0.17	-0.11
Symbol cancelation	0.18	0.67	0.23					
Digit Span, Backwards score	0.85	0.85	0.53					
Digit Span, Forwards score	0.70	0.70	0.58					
Language/fluency								
Naming common objects	0.63	0.63	-2.19	-1.01				
Animal naming	0.09	0.48	0.46					
Writing or saying sentence	0.58	0.58	-1.13					
Repeat a phrase	0.64	0.64	-1.18					
Close your eyes	0.66	0.66	-1.28	0.16	0.90			
3 stage task	0.60	0.60	-1.97	-1.38	-0.56			
Naming described objects	0.70	0.70	-2.47	-2.10	-1.58	-1.06	-0.50	0.28
Visuospatial								
Interlocking pentagons	0.88	0.88	0.59	0.71				
Constructional praxis, immediate	0.88	0.88	-1.34	-0.35	0.34	1.28		

Supplemental Table 2. Fit statistics for unidimensional and hierarchical confirmatory factor analysis models among participants without dementia: Results from LASI-DAD (N=733)

Domain specificity	Model	Number of items	RMSEA	CFI	SRMR	Descriptor
	Single domain models					
Broad	Orientation	11	0.061	0.956	0.091	Good
Narrow	Memory-Episodic-Immediate	4	0.021	0.999	0.008	Good
Narrow	Memory-Episodic-Delayed	5	0.041	0.991	0.015	Good
Narrow	Memory-Episodic-Recognition	2	0.000	1.000	0.000	Perfect
Narrow	EF-Abstract Reasoning	4	0.033	0.998	0.005	Adequate
Narrow	EF-Attention/Speed	5	0.076	0.983	0.028	Adequate
Broad	Language/fluency	7	0.040	0.976	0.033	Good
Broad	Visuospatial	2	0.000	1.000	0.000	Perfect
	Multiple domain model					
	Full hierarchical CFA	40	0.045	0.928	0.071	Adequate

Supplemental information 1: Imputation of cognitive data

Table 1 shows the percentage of observations with missing data on each cognitive test; no variable was missing more than 15% in the sample, so as is common in survey data, we *imputed* most missing observations. The goal of imputation is to replace the missing values with random draws from a conditional distribution such that the estimated joint distribution from the completed (imputed) data is an unbiased estimator of the true joint distribution of these variables (1-2).

For the most part, we imputed missing cognitive test variables studied in this paper for participants for whom the tests were missing. However, some variables were only administered in specific samples, such as among literate or illiterate respondents. We did not impute scores for the samples that tests were not administered to. For cognitive test items, we recoded “don't know” as incorrect. There are some indications that other missingness, especially refusals, may also sometimes indicate that the respondent does not know the correct answer, but because we cannot be sure about this, we treated refusals as missing scores to be imputed. Additionally, we imputed the serial 7s score for individuals who cannot count, even though strictly speaking the individual gave no correct answers and would not be able to complete the task. However, we use serial 7s as a test of processing speed and attention and not as a test of numeric ability. Hence, a score of 0 for individuals who were not administered the test because they cannot count would not necessarily reflect their processing speed and attention well.

The imputation method we implemented was inspired by the imputations of cognition variables in the HRS (3). It is also similar to the method used in the Survey of Health, Ageing and Retirement in Europe (SHARE) (4). We specified a regression model for each cognition variable as a function of the other cognition variables and a rich set of background variables: health, demographics, and socio-economic variables, as well as reports from an informant about the individual's cognition. The regression model specifies the conditional distribution of the variable that must be imputed as a function of the regressors, and the imputations are pseudo-random draws from this conditional distribution. Take, for example, a binary variable such as whether the respondent correctly answered the question about what year it is. Let this variable be *y* and let the regressors be collected in the vector **x**. We specified a logistic regression model for *y* as a function of **x**:

$$\Pr(y_i = 1|\mathbf{x}_i) = p_i = \frac{e^{\mathbf{x}_i'\beta}}{1+e^{\mathbf{x}_i'\beta}}$$

This was estimated in the sample where y_i is observed. Then, we generated a pseudo-random draw u_i from a uniform distribution on the interval (0,1) and for the sample where y_i was missing, we computed p_i and imputed $y_i = 1$ if $u_i \leq p_i$ and $y_i = 0$ otherwise. For binary variables, we used (binary) logistic regression (i.e., logit) models; for ordinal variables, we used ordered logit; for count variables, we used negative binomial regression; and for unordered categorical variables, we used multinomial logit.

Because respondents can have multiple missing variables, one or more of the regressors in **x** could themselves be missing. These needed to be imputed first. As in HRS and SHARE, we used chained imputation (also known as fully conditional specification) (5-6). This cycles over the cognition variables, in which each of them is imputed in turn, with the other cognition variables and background variables as regressors, and then repeats this cycle multiple times. We used one cycle to initialize the chain and 10 cycles (iterations) to update the imputations. Cycling over all cognition variables and background variables at the same time would be rather daunting, and therefore, we followed the HRS's lead and imputed blocks of variables sequentially; cognitive variables were imputed after imputation of time-varying and time-invariant demographics and health variables. We also used imputations of income and wealth from the core survey, categorized into quintiles, as regressors for the health and cognition variables.

In many cases, we did not use all items of a scale separately as regressors, but used summary scores (aggregates) instead. For example, when imputing a word recall item, we used the number correct in the orientation to place scale (0-5) instead of all five items separately. An exception is that when imputing an item within a scale, we used the other items separately, so when imputing the “city” variable, we used the state, name, address, and floor variables separately as regressors.

Despite all these aggregations and simplifications, the number of regressors in the cognition models was very large (about 100), which sometimes caused numerical problems in estimation of an imputation model. If this happened, we first dropped a few variables that often caused problems because of collinearities, and if this did not solve the problem, we used a smaller set of high-level aggregates. These high-level aggregates were obtained by grouping the regressors in a few sets (socio-economic status, physical health, mental health, activities, and five sets of cognition variables) and using the first principal component of each set, in addition to gender and age group.

Supplemental References

1. Little, R. J. A., & Rubin, D. B. (2002). Statistical analysis with missing data (2nd ed.). New York, NY: Wiley.
2. Lee, J., Meijer, E., & Phillips, D. (2015). The effect of using different imputation methods for economic variables in aging surveys. Working Paper No. 2015-019. Los Angeles, CA: University of Southern California, Center for Economic and Social Research.
3. Fisher, G. G., Hassan, H., Faul, J. D., Rodgers, W. L., & Weir, D. R. (2017). Health and Retirement Study: Imputation of Cognitive Functioning Measures: 1992 – 2014 (Final Release Version): Data Description. Ann Arbor, MI: University of Michigan, Survey Research Center.
4. De Luca, G., Celidoni, M., & Trevisan, E. (2015). Item nonresponse and imputation strategies in SHARE Wave 5. In F. Malter & A. Börsch-Supan (Eds.), SHARE Wave 5: Innovations & Methodology (pp. 85-100). Munich: MEA.
5. Raghunathan, T. E., Lepkowski, J. M., van Hoewyk, J., & Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. Survey Methodology, 27, 85–95.
6. van Buuren, S., Brand, J. P. L., Groothuis-Oudshoorn, C. G. M., & Rubin, D. B. (2006). Fully conditional specification in multivariate imputation. Journal of Statistical Computation and Simulation, 76, 1049–1064.