



THE UNIVERSITY OF  
CHICAGO

# A model for cognitive function assessments in surveys incorporating response times

---

Phil Schumm

February 23, 2023

Department of Public Health Sciences

Part of project led by Seth Sanders (R21HD088618), with support from the National Social Life, Health and Aging Project (R01AG043538, Linda Waite PI)

1. Background
2. Approach
  - Statistical Model
  - Estimation in Stata using `gsem`
3. Results
4. Future Directions

# Background

---

# Montreal Cognitive Assessment (MoCA)

- Screening tool for detecting mild cognitive impairment (MCI)
- Assesses (1) orientation, (2) memory recall, (3) visuospatial ability, (3) executive function, (4) attention, concentration and working memory, (5) language, and (6) abstract reasoning.
- Administered in the clinic setting in approximately 10 minutes
- Scored by adding number of items correct (0–30)
- Has been translated into multiple languages

# National Social Life, Health and Aging Project (NSHAP)

- NSHAP is a national, longitudinal study of health and aging started in 2005
- Probability sample of adults aged 50 and older
- Primary objective is to study association between social factors and differences in health trajectories at older ages
- In-home interviews in 2005, 2010 and 2015
- Collects wide range of social measures, as well as objective (physical measures and biosamples) and self-report measures of health
- NSHAP developed a survey-adapted version of the MoCA (MoCA-SA) in 2010
- Data are available from the National Archive of Computerized Data on Aging (NACDA)

Question 1: Does the time required to complete parts of the survey reflect cognitive function?

- CAPI instrument included time stamps between each survey item
- These represent the *total time* to ask the question, answer it and record that answer (not only *response time*)
- Time taken to answer MoCA-SA items correlates with MoCA-SA score, change in score over time and 5-year mortality<sup>1</sup>

---

<sup>1</sup> “Using Response Time Data from Social Science Surveys to Model Cognition and Cognitive Decline” by Seth Sanders, William Dale, Henrik Olsson, Lynne Schofield, Phil Schumm and Linda Waite, 2017.

Question 1: Does the time required to complete parts of the survey reflect cognitive function?

- CAPI instrument included time stamps between each survey item
- These represent the *total time* to ask the question, answer it and record that answer (not only *response time*)
- Time taken to answer MoCA-SA items correlates with MoCA-SA score, change in score over time and 5-year mortality<sup>1</sup>

Question 2: Can time taken to complete items be used to augment existing measures of cognitive function?

---

<sup>1</sup> "Using Response Time Data from Social Science Surveys to Model Cognition and Cognitive Decline" by Seth Sanders, William Dale, Henrik Olsson, Lynne Schofield, Phil Schumm and Linda Waite, 2017.

# Approach

---



## Two-parameter logistic item response model

Let  $y_{ij}$  be a binary (0/1) variable indicating whether the  $i$ th respondent correctly answered the  $j$ th item:

$$\text{logit}[P(y_{ij} = 1|c_i)] = \alpha_j + \beta_{1j}c_i \quad (1)$$

$-\alpha_j$  are known as the item difficulties

$c_i \sim \mathcal{N}(0, 1)$  represent each respondent's overall cognitive function

$\beta_{1j}$  are referred to as discrimination parameters (or factor loadings)

When cognitive function itself is viewed as the outcome, we may further specify a mean model for the  $c_i$  such as  $c_i \sim \mathcal{N}(\mathbf{x}_i'\boldsymbol{\beta}, 1)$ .

When cognitive function takes the role of a predictor, we may use estimates  $\hat{c}_i$  such as the empirical Bayes means.

## Incorporating Response Times

To incorporate response times, we augment Model 1 by permitting the time taken by respondent  $i$  to answer the  $j$ th item ( $t_{ij}$ ) to affect the probability of a correct response:

$$\text{logit}[P(y_{ij} = 1|c_i, t_{ij})] = \alpha_j + \beta_{1j}c_i + \beta_{2j}f_j(\log(t_{ij})) \quad (2)$$

While estimates  $\hat{\beta}_{2j}$  and  $\hat{f}_j(\cdot)$  from Equation 2 are informative with respect to the association between an item's response time and its success probability, simply adding  $t_{ij}$  to this equation does not improve our estimates of  $c_i$ . This is because estimation of Equation 2 alone typically involves assuming that  $c_i$  and  $f_j(\log(t_{ij}))$  are uncorrelated. Thus, to the extent that longer response times may reflect a decrease in cognitive function, this will not be captured in our estimates of  $c_i$ .

## Incorporating Response Times

To incorporate response times, we augment Model 1 by permitting the time taken by respondent  $i$  to answer the  $j$ th item ( $t_{ij}$ ) to affect the probability of a correct response:

$$\text{logit}[P(y_{ij} = 1 | c_i, t_{ij})] = \alpha_j + \beta_{1j}c_i + \beta_{2j}f_j(\log(t_{ij})) \quad (2)$$

To address this, we jointly estimate a second equation predicting the response times:

$$\log(t_{ijk}) = \alpha_j^* + \beta_{1j}^*c_i + \beta_{2j}^*d_k + \mathbf{z}'_{ij}\beta_3^* + \epsilon_{ijk} \quad (3)$$

- Effect of cognitive function on completion time ( $\beta_{1j}^*$ ) allowed to vary across items
- Interviewer effects  $d_k \sim \mathcal{N}(0, \sigma^2)$
- Respondent-level covariates  $\mathbf{z}_{ij}$  (e.g., account for specific disability)

# Estimating in Stata using `gsem` (example with 3 items only)

```
. des fi_id-logt_digits5
```

Variable name	Storage type	Display format	Value label	Variable label
fi_id	str6	%6s		Interviewer ID
moca_contour	byte	%9.0g		Clock score (contour)
moca_numbers	byte	%9.0g		Clock score (numbers)
moca_hands	byte	%9.0g		Clock score (hands)
moca_trail2	byte	%9.0g		Trails score
moca_5numbers	byte	%9.0g		Forward digits score
logt_clock	float	%9.0g		Log time (seconds) for clock draw
logt_trail	float	%9.0g		Log time (seconds) for trails
logt_digits5	float	%9.0g		Log time (seconds) for forward digits

## Estimating in Stata using `gsem` (example with 3 items only)

```
mkspline f_clock_s = logt_clock, cubic
mkspline f_trail_s = logt_trail, cubic
mkspline f_digits5_s = logt_digits5, cubic

gsem (moca_contour <- f_clock_s* C, logit) ///
     (moca_numbers <- f_clock_s* C, logit) ///
     (moca_hands <- f_clock_s* C, logit) ///
     (logt_clock <- C I[fi_id]) ///
     (moca_trail2 <- f_trail_s* C, logit) ///
     (logt_trail <- C I[fi_id]) ///
     (moca_5numbers <- f_digits5_s* C, logit) ///
     (logt_digits5 <- C I[fi_id]), var(C@1) vce(robust)
```

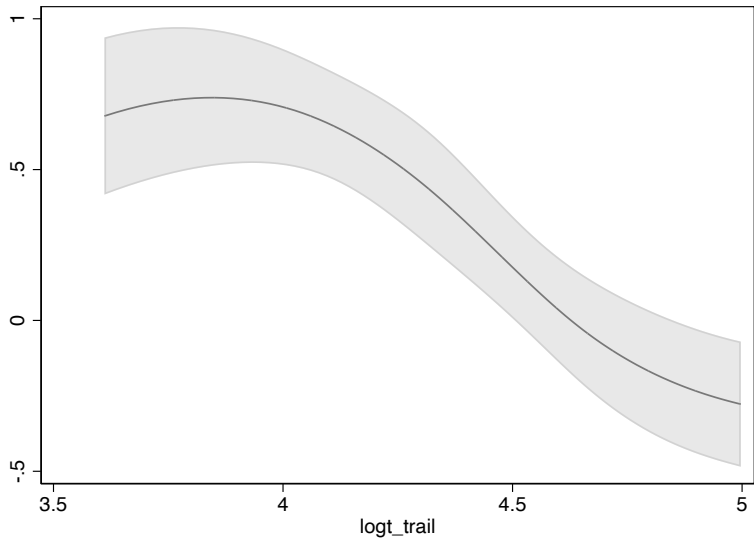
# Results

---

## Results: Trails

-----+-----							
moca_trail2							
f_trail_s1		1.097339	.1676696	6.54	0.000	.7687125	1.425965
f_trail_s2		-2.479554	.6524467	-3.80	0.000	-3.758326	-1.200782
f_trail_s3		3.790759	10.41773	0.36	0.716	-16.62762	24.20914
f_trail_s4		15.39633	21.19677	0.73	0.468	-26.14858	56.94124
C		1.013618	.1086819	9.33	0.000	.8006056	1.226631
_cons		-3.151297	.5333989	-5.91	0.000	-4.196739	-2.105854
-----+-----							
logt_trail							
I[fi_id]		.9560522	.1512955	6.32	0.000	.6595184	1.252586
C		-.14542	.0399833	-3.64	0.000	-.2237858	-.0670543
_cons		4.284304	.029731	144.10	0.000	4.226032	4.342576
-----+-----							

## Results: Trails

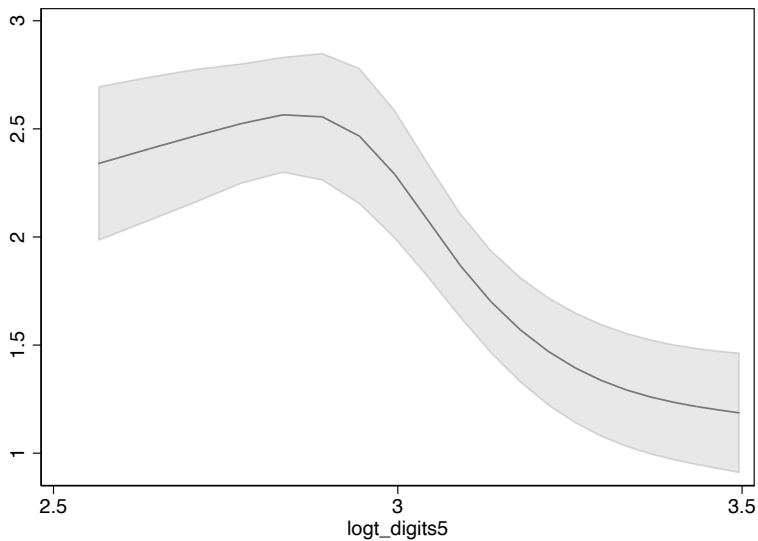




## Results: Forward digits

-----+-----							
moca_5numbers							
f_digits5_s1		.9400005	.6979078	1.35	0.178	-.4278737	2.307875
f_digits5_s2		-.4494738	4.65836	-0.10	0.923	-9.579691	8.680743
f_digits5_s3		-37.80609	26.02594	-1.45	0.146	-88.81599	13.20381
f_digits5_s4		105.262	42.70536	2.46	0.014	21.56103	188.963
C		.6644854	.0959759	6.92	0.000	.4763761	.8525947
_cons		-.0715235	1.840264	-0.04	0.969	-3.678374	3.535327
-----+-----							
logt_digits5							
I[fi_id]		.9565751	.2007805	4.76	0.000	.5630526	1.350098
C		-.0790521	.0099747	-7.93	0.000	-.0986021	-.0595021
_cons		2.953775	.0204893	144.16	0.000	2.913616	2.993933
-----+-----							

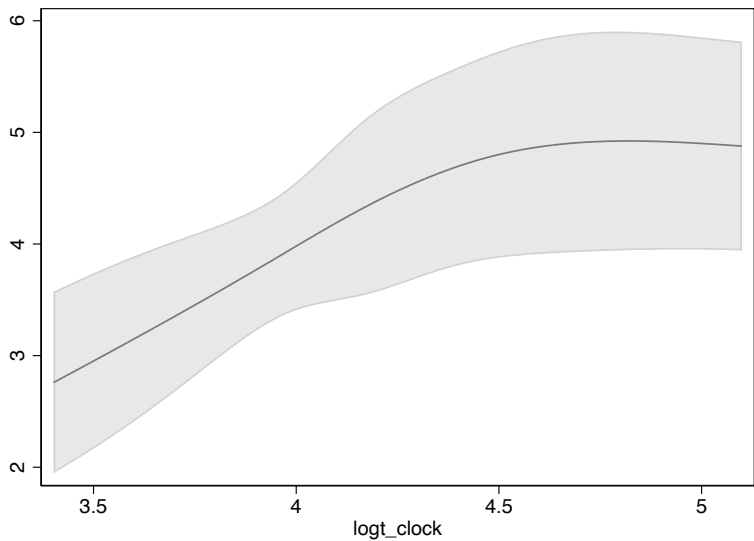
## Results: Forward digits



# Results: Clock

-----							
moca_contour							
f_clock_s1		1.902008	.350201	5.43	0.000	1.215627	2.588389
f_clock_s2		.4251896	2.49048	0.17	0.864	-4.456061	5.30644
f_clock_s3		-9.347134	39.50107	-0.24	0.813	-86.7678	68.07354
f_clock_s4		5.857761	66.55008	0.09	0.930	-124.578	136.2935
C		1.44437	.2783021	5.19	0.000	.8989076	1.989832
_cons		-3.715171	1.164329	-3.19	0.001	-5.997214	-1.433127
-----							
moca_numbers							
f_clock_s1		1.61253	.3413945	4.72	0.000	.9434088	2.281651
f_clock_s2		-.75063	.9789638	-0.77	0.443	-2.669364	1.168104
f_clock_s3		-33.21954	15.2335	-2.18	0.029	-63.07664	-3.362439
f_clock_s4		78.09284	26.95085	2.90	0.004	25.27015	130.9155
C		1.575744	.1587984	9.92	0.000	1.264505	1.886983
_cons		-4.576671	1.256111	-3.64	0.000	-7.038604	-2.114738
-----							
moca_hands							
f_clock_s1		1.256381	.3333291	3.77	0.000	.6030684	1.909694
f_clock_s2		-.802166	.9002886	-0.89	0.373	-2.566699	.9623672
f_clock_s3		-22.96024	14.09194	-1.63	0.103	-50.57993	4.659449
f_clock_s4		55.69789	25.12771	2.22	0.027	6.448484	104.9473
C		1.572536	.1452214	10.83	0.000	1.287907	1.857164
_cons		-4.669391	1.237165	-3.77	0.000	-7.094191	-2.244591
-----							
logt_clock							
I[fi_id]		1	(constrained)				
C		-.2780116	.0388656	-7.15	0.000	-.3541869	-.2018364
_cons		4.169707	.0281429	148.16	0.000	4.114548	4.224867
-----							

## Results: Clock contour



## Results: Interviewer effects

```
-----+-----  
var(I[fi_id])| .0465387 .017645 .0221352 .0978462  
var(C)| 1 (constrained)  
-----+-----
```

## **Future Directions**

---

## Future Directions

- Augmenting IRT model
  - Guessing parameter
  - Bi-factor model to account for additional correlation among items
- Evaluating improvement in reliability of cognitive function measurements due to incorporation of time
- Incorporating completion times from other parts of the survey
- Using response times for remote data collection (e.g., web or phone)