CODA Cohort Random Effects Cognition Modeling

Zach Kunicki

## Background and purpose of study

This report describes the data management and results of modeling cognition in the Children of the Depression (CODA) cohort of the Health and Retirement Study (HRS).

The purpose of this study is to study the concept of normative cognitive aging, which for the purpose of this study is defined as aging-related changes in the absence of a diagnosis of a neurocognitive disorder. This is a major unanswered question in the field of cognitive aging, which the CODA cohort is uniquely suited to answer. First, CODA is a community sample which reduces concerns of selection bias when compared to samples recruited from a specialty setting. Second, the CODA cohort began recruitment in 1998, with follow-up surveys currently on going. Data until 2020 are currently available, for 22 years of follow up assessments. Third, there is age homogeneity as the entire sample was between 68-74 at baseline. This also reduces selection bias, as samples with age heterogeneity (e.g., 68-85) cannot be certain that younger participants are representative of older participants (i.e., the older participants were healthy enough to make it to their current age, and the younger participants may not live that long). Fourth, it is possible to adjust for practice and retest effects that are commonly seen in cognitive aging cohorts.

A fifth ideal component of CODA is the ability to use both a cognitive composite score and individual cognitive tests to examine change in cognition over time. Composite scores and individual cognitive tests each have advantages and disadvantages. Epidemiologists may prefer cognitive comoposites because they are a single, normally distributed variable that is easy to work with in analytical models. Neuropsychologists may prefer individual cognitive tests because they can give information on domain-specific functioning that could be useful for diagnosis or treatment.

The purpose of the current study is to develop a cognitive composite score for use in the CODA cohort, which can then be used to examine normative cognitive aging. As part of this analysis, we will consider differential item functioning as there are known mode of assessment effects in the HRS (see Smith et al., 2022 and Domingue et al., 2023).

## Step 1: Selecting cases to include

The first step of the analysis was selecting members of the CODA cohort. In the HRS tracker file, we filtered for participants who had values of 21 for STUDY. STUDY is defined as “Study membership is based on the entry cohort of the individual. This will differ from the birth cohort for age-ineligible spouses or partners” and 21 is defined as “21. CODA (Children of the Depression Age).”

We identified N = 2464 in the overall cohort. We dropped N = 6 for having a 1998 age outside of the CODA-eligible age band of 68-74. There were also N = 163 participants with a weight of 0, who were excluded from the analysis. A weight of 0 means the participant was ineligible for the study. Reasons for having a weight of zero included “not cohort eligible this wave,” “nursing home resident”, and “xWGTR greater than zero or xIWTYPE greater than one.” xWGTR is defined as “the sampling weight for analysis at the respondent level for respondents living in the community” (HRS Tracker File, 2023, pg. 21). xIWTYPE refers a 7 category variable about the type of interview conducted. A value of 1 is the core interview, values of 5, 11, 15, 21, 25, and 99 refer to core interview not obtained, exit interview obtained, exit interview not obtained, post-exit interview obtained, post-exit interview not obtained, and not in the sample at this wave (HRS Tracker File, 2023, pg. 25). *Table 1* shows the break down for why a participant had a weight of 0 and was removed from the primary analyses.

### Table 1: Summary of participants excluded for having a weight of zero

| **Characteristic** | **N = 159** |
| --- | --- |
| FWHY0WGT |  |
| Not cohort eligible this wave | 17 (11%) |
| Nursing home resident | 2 (1.3%) |
| xWGTR greater than zero or xIWTYPE greater than one | 140 (88%) |
| FIWTYPE |  |
| Core interview obtained | 19 (12%) |
| Core interview not obtained | 70 (44%) |
| Not in the sample this wave | 70 (44%) |

Note that for the participants listed as “Core interview obtained” under FIWTYPE, these same participants either were not cohort eligible or a nursing home resident under FWHY0WGT, as shown in Table 2.

### Table 2: Breakdown of FIWTYPE and FWHY0WGT

| FWHY0WGT | FIWTYPE | n |
| --- | --- | --- |
| Not cohort eligible this wave | Core interview obtained | 17 |
| Nursing home resident | Core interview obtained | 2 |
| xWGTR greater than zero or xIWTYPE greater than one | Core interview not obtained | 70 |
| xWGTR greater than zero or xIWTYPE greater than one | Not in the sample this wave | 70 |

## Step 2: Scoring variables

There are six cognitive variables in the HRS cognitive battery: immediate recall, delayed recall, backwards counting, object naming, orientation to time, and serial 7s. I will describe those tests in an upcoming revision (or ask a real neuropsychologist to do it).

McArdle et al. (2007) scored these variables on a 0-100% scale. In my code, I recreate that scoring but the final analytical variables are done on a conventional 0, 1, 2… scale. We also do not use the immediate recall variable because immediate recall is a cognitive task that will have a very strong ceiling effect until the later stages of dementia, and as such is not a good indicator of cognition in community-based samples.

Table 3 shows the scored cognitive variables. Here, we see that there are very sparse entries at the lower end for object naming and orientation to time.

### Table 3: Scored cognitive variables

| **Characteristic** | **0**, N = 2,295 | **2**, N = 2,060 | **4**, N = 1,872 | **6**, N = 1,698 | **8**, N = 1,537 | **10**, N = 1,373 | **12**, N = 1,164 | **14**, N = 1,033 | **16**, N = 828 | **18**, N = 604 | **20**, N = 405 | **22**, N = 263 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| vddr | 3.82 (2.22) | 3.69 (2.26) | 4.05 (2.08) | 3.80 (1.97) | 3.56 (1.97) | 3.47 (1.97) | 3.19 (1.94) | 2.89 (2.01) | 2.83 (1.99) | 2.70 (1.92) | 2.80 (1.90) | 2.66 (1.83) |
| Unknown | 0 | 0 | 189 | 164 | 136 | 103 | 128 | 126 | 108 | 91 | 62 | 80 |
| vdbc | 2,055 (96%) | 1,814 (96%) | 1,619 (96%) | 1,466 (95%) | 1,334 (94%) | 1,178 (93%) | 939 (91%) | 818 (91%) | 631 (88%) | 459 (90%) | 313 (91%) | 169 (88%) |
| Unknown | 147 | 171 | 181 | 152 | 115 | 109 | 130 | 131 | 114 | 93 | 61 | 70 |
| vdda |  |  |  |  |  |  |  |  |  |  |  |  |
| 0 | 3 (0.1%) | 3 (0.2%) | 2 (0.1%) | 6 (0.4%) | 10 (0.7%) | 7 (0.6%) | 4 (0.4%) | 7 (0.8%) | 8 (1.2%) | 3 (0.6%) | 3 (1.0%) | 2 (1.1%) |
| 1 | 10 (0.5%) | 7 (0.4%) | 6 (0.4%) | 7 (0.5%) | 9 (0.6%) | 11 (0.9%) | 14 (1.4%) | 15 (1.8%) | 11 (1.7%) | 5 (1.1%) | 9 (2.9%) | 6 (3.4%) |
| 2 | 44 (2.0%) | 24 (1.3%) | 35 (2.1%) | 31 (2.0%) | 47 (3.4%) | 39 (3.1%) | 48 (4.9%) | 44 (5.1%) | 32 (4.8%) | 28 (6.0%) | 13 (4.2%) | 14 (8.0%) |
| 3 | 345 (16%) | 308 (16%) | 241 (14%) | 242 (16%) | 251 (18%) | 219 (18%) | 209 (21%) | 211 (25%) | 146 (22%) | 104 (22%) | 90 (29%) | 37 (21%) |
| 4 | 1,759 (81%) | 1,568 (82%) | 1,402 (83%) | 1,260 (82%) | 1,082 (77%) | 963 (78%) | 710 (72%) | 579 (68%) | 466 (70%) | 328 (70%) | 193 (63%) | 117 (66%) |
| Unknown | 134 | 150 | 186 | 152 | 138 | 134 | 179 | 177 | 165 | 136 | 97 | 87 |
| vdnm |  |  |  |  |  |  |  |  |  |  |  |  |
| 0 | 1 (<0.1%) | 1 (<0.1%) | 0 (0%) | 0 (0%) | 3 (0.3%) | 0 (0%) | 1 (0.2%) | 1 (0.2%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) |
| 1 | 10 (0.5%) | 7 (0.4%) | 1 (<0.1%) | 1 (<0.1%) | 0 (0%) | 1 (<0.1%) | 3 (0.5%) | 3 (0.5%) | 2 (0.5%) | 0 (0%) | 0 (0%) | 1 (0.9%) |
| 2 | 65 (3.1%) | 25 (1.4%) | 9 (0.6%) | 10 (0.7%) | 6 (0.5%) | 7 (0.7%) | 10 (1.6%) | 11 (1.9%) | 7 (1.6%) | 6 (2.0%) | 1 (0.6%) | 4 (3.4%) |
| 3 | 241 (12%) | 142 (7.7%) | 198 (14%) | 135 (10%) | 110 (9.3%) | 93 (9.0%) | 119 (19%) | 80 (14%) | 65 (15%) | 33 (11%) | 38 (23%) | 18 (16%) |
| 4 | 1,768 (85%) | 1,665 (90%) | 1,201 (85%) | 1,200 (89%) | 1,062 (90%) | 928 (90%) | 485 (78%) | 496 (84%) | 358 (83%) | 266 (87%) | 129 (77%) | 93 (80%) |
| Unknown | 210 | 220 | 463 | 352 | 356 | 344 | 546 | 442 | 396 | 299 | 237 | 147 |
| vds7 |  |  |  |  |  |  |  |  |  |  |  |  |
| 0 | 76 (4.4%) | 59 (3.9%) | 11 (0.8%) | 12 (0.9%) | 12 (1.0%) | 15 (1.5%) | 15 (1.9%) | 18 (2.6%) | 11 (2.0%) | 6 (1.6%) | 7 (2.6%) | 2 (1.3%) |
| 1 | 290 (17%) | 254 (17%) | 37 (2.7%) | 40 (3.1%) | 29 (2.5%) | 35 (3.5%) | 25 (3.1%) | 32 (4.6%) | 34 (6.1%) | 17 (4.5%) | 18 (6.6%) | 6 (3.9%) |
| 2 | 233 (13%) | 211 (14%) | 95 (6.8%) | 100 (7.9%) | 76 (6.6%) | 69 (6.9%) | 69 (8.6%) | 56 (8.0%) | 54 (9.7%) | 41 (11%) | 23 (8.5%) | 20 (13%) |
| 3 | 85 (4.9%) | 71 (4.7%) | 183 (13%) | 198 (16%) | 171 (15%) | 155 (16%) | 156 (20%) | 128 (18%) | 103 (19%) | 73 (19%) | 47 (17%) | 19 (12%) |
| 4 | 143 (8.3%) | 104 (6.9%) | 313 (23%) | 289 (23%) | 289 (25%) | 240 (24%) | 183 (23%) | 183 (26%) | 143 (26%) | 101 (27%) | 70 (26%) | 46 (30%) |
| 5 | 906 (52%) | 801 (53%) | 752 (54%) | 632 (50%) | 574 (50%) | 481 (48%) | 352 (44%) | 283 (40%) | 210 (38%) | 141 (37%) | 106 (39%) | 60 (39%) |
| Unknown | 562 | 560 | 481 | 427 | 386 | 378 | 364 | 333 | 273 | 225 | 134 | 110 |

Based on this, we re-score both variables. In orientation to time, we collapse the bottom two categories and for object naming we collapse the bottom three categories. The rescoring is shown in Tables 4 and 5, respectively. An updated table with all of the scored variables is shown in Table 6.

### Table 4: Rescoring of orientation time

| vdda | vdda\_r | n |
| --- | --- | --- |
| 0 | 0 | 58 |
| 1 | 0 | 110 |
| 2 | 1 | 399 |
| 3 | 2 | 2403 |
| 4 | 3 | 10427 |
| NA | NA | 1735 |

### Table 5: Rescoring of object naming

| vdnm | vdnm\_r | n |
| --- | --- | --- |
| 0 | 0 | 7 |
| 1 | 0 | 29 |
| 2 | 0 | 161 |
| 3 | 1 | 1272 |
| 4 | 2 | 9651 |
| NA | NA | 4012 |

### Table 6: Rescored cognitive variables

| **Characteristic** | **0**, N = 2,295 | **2**, N = 2,060 | **4**, N = 1,872 | **6**, N = 1,698 | **8**, N = 1,537 | **10**, N = 1,373 | **12**, N = 1,164 | **14**, N = 1,033 | **16**, N = 828 | **18**, N = 604 | **20**, N = 405 | **22**, N = 263 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| vddr | 3.82 (2.22) | 3.69 (2.26) | 4.05 (2.08) | 3.80 (1.97) | 3.56 (1.97) | 3.47 (1.97) | 3.19 (1.94) | 2.89 (2.01) | 2.83 (1.99) | 2.70 (1.92) | 2.80 (1.90) | 2.66 (1.83) |
| Unknown | 0 | 0 | 189 | 164 | 136 | 103 | 128 | 126 | 108 | 91 | 62 | 80 |
| vdbc | 2,055 (96%) | 1,814 (96%) | 1,619 (96%) | 1,466 (95%) | 1,334 (94%) | 1,178 (93%) | 939 (91%) | 818 (91%) | 631 (88%) | 459 (90%) | 313 (91%) | 169 (88%) |
| Unknown | 147 | 171 | 181 | 152 | 115 | 109 | 130 | 131 | 114 | 93 | 61 | 70 |
| vdda\_r |  |  |  |  |  |  |  |  |  |  |  |  |
| 0 | 13 (0.6%) | 10 (0.5%) | 8 (0.5%) | 13 (0.8%) | 19 (1.4%) | 18 (1.5%) | 18 (1.8%) | 22 (2.6%) | 19 (2.9%) | 8 (1.7%) | 12 (3.9%) | 8 (4.5%) |
| 1 | 44 (2.0%) | 24 (1.3%) | 35 (2.1%) | 31 (2.0%) | 47 (3.4%) | 39 (3.1%) | 48 (4.9%) | 44 (5.1%) | 32 (4.8%) | 28 (6.0%) | 13 (4.2%) | 14 (8.0%) |
| 2 | 345 (16%) | 308 (16%) | 241 (14%) | 242 (16%) | 251 (18%) | 219 (18%) | 209 (21%) | 211 (25%) | 146 (22%) | 104 (22%) | 90 (29%) | 37 (21%) |
| 3 | 1,759 (81%) | 1,568 (82%) | 1,402 (83%) | 1,260 (82%) | 1,082 (77%) | 963 (78%) | 710 (72%) | 579 (68%) | 466 (70%) | 328 (70%) | 193 (63%) | 117 (66%) |
| Unknown | 134 | 150 | 186 | 152 | 138 | 134 | 179 | 177 | 165 | 136 | 97 | 87 |
| vdnm\_r |  |  |  |  |  |  |  |  |  |  |  |  |
| 0 | 76 (3.6%) | 33 (1.8%) | 10 (0.7%) | 11 (0.8%) | 9 (0.8%) | 8 (0.8%) | 14 (2.3%) | 15 (2.5%) | 9 (2.1%) | 6 (2.0%) | 1 (0.6%) | 5 (4.3%) |
| 1 | 241 (12%) | 142 (7.7%) | 198 (14%) | 135 (10%) | 110 (9.3%) | 93 (9.0%) | 119 (19%) | 80 (14%) | 65 (15%) | 33 (11%) | 38 (23%) | 18 (16%) |
| 2 | 1,768 (85%) | 1,665 (90%) | 1,201 (85%) | 1,200 (89%) | 1,062 (90%) | 928 (90%) | 485 (78%) | 496 (84%) | 358 (83%) | 266 (87%) | 129 (77%) | 93 (80%) |
| Unknown | 210 | 220 | 463 | 352 | 356 | 344 | 546 | 442 | 396 | 299 | 237 | 147 |
| vds7 |  |  |  |  |  |  |  |  |  |  |  |  |
| 0 | 76 (4.4%) | 59 (3.9%) | 11 (0.8%) | 12 (0.9%) | 12 (1.0%) | 15 (1.5%) | 15 (1.9%) | 18 (2.6%) | 11 (2.0%) | 6 (1.6%) | 7 (2.6%) | 2 (1.3%) |
| 1 | 290 (17%) | 254 (17%) | 37 (2.7%) | 40 (3.1%) | 29 (2.5%) | 35 (3.5%) | 25 (3.1%) | 32 (4.6%) | 34 (6.1%) | 17 (4.5%) | 18 (6.6%) | 6 (3.9%) |
| 2 | 233 (13%) | 211 (14%) | 95 (6.8%) | 100 (7.9%) | 76 (6.6%) | 69 (6.9%) | 69 (8.6%) | 56 (8.0%) | 54 (9.7%) | 41 (11%) | 23 (8.5%) | 20 (13%) |
| 3 | 85 (4.9%) | 71 (4.7%) | 183 (13%) | 198 (16%) | 171 (15%) | 155 (16%) | 156 (20%) | 128 (18%) | 103 (19%) | 73 (19%) | 47 (17%) | 19 (12%) |
| 4 | 143 (8.3%) | 104 (6.9%) | 313 (23%) | 289 (23%) | 289 (25%) | 240 (24%) | 183 (23%) | 183 (26%) | 143 (26%) | 101 (27%) | 70 (26%) | 46 (30%) |
| 5 | 906 (52%) | 801 (53%) | 752 (54%) | 632 (50%) | 574 (50%) | 481 (48%) | 352 (44%) | 283 (40%) | 210 (38%) | 141 (37%) | 106 (39%) | 60 (39%) |
| Unknown | 562 | 560 | 481 | 427 | 386 | 378 | 364 | 333 | 273 | 225 | 134 | 110 |

Finally, we obtain the demographic variables that will be used in the final analysis. These are shown in Table 4. Note that these demographics are all based on the **unweighted** sample. In the final publication version, we will report the *unweighted* N but the *weighted* statistics in terms of Means/SDs and percentages. This table is shown in Table 8.

### Table 7: Unweighted Demographics

| **Characteristic** | **N = 2,295** |
| --- | --- |
| vdfem | 1,359 (59%) |
| vdhisp | 133 (5.8%) |
| Unknown | 1 |
| vdothrac | 101 (4.4%) |
| vdblack | 209 (9.1%) |
| NUMBER OF YEARS IN SCHOOL | 11.9 (3.4) |
| AGE AT 1998 INTERVIEW |  |
| 67 | 126 (5.5%) |
| 68 | 321 (14%) |
| 69 | 314 (14%) |
| 70 | 375 (16%) |
| 71 | 330 (14%) |
| 72 | 327 (14%) |
| 73 | 325 (14%) |
| 74 | 177 (7.7%) |

### Table 8: Weighted Demographics

| **Characteristic** | **N = 2295** |
| --- | --- |
| vdfem | 61.6% |
| vdhisp | 5.4% |
| vdrace |  |
| Black | 10.0% |
| Other race | 4.5% |
| White | 85.5% |
| NUMBER OF YEARS IN SCHOOL | 11.80 (3.34) |
| AGE AT 1998 INTERVIEW |  |
| 67 | 5.5% |
| 68 | 14.4% |
| 69 | 13.4% |
| 70 | 16.2% |
| 71 | 14.1% |
| 72 | 14.4% |
| 73 | 14.4% |
| 74 | 7.7% |

## Step 3: Model cognition at baseline

The first preliminary model examines cognition at baseline. In this model, we fix the factor variance at 1 and our goal is to estimate the first loading for use in subsequent models. The model syntax is:

### Model 1 Syntax

MODEL = "GCP BY vddr\* vdbc vdda vdnm vds7;  
 GCP@1;"