SUPPLEMENTARY MATERIALS TO SCHAEFER ET AL.,

IS LOW COGNITIVE FUNCTIONING A PREDICTOR OR CONSEQUENCE OF MAJOR DEPRESSIVE DISORDER? A TEST IN TWO LONGITUDINAL BIRTH COHORTS

These Supplementary Materials contain additional details about the neuropsychological tests administered at ages 13 and 38 in the Dunedin Study. In addition, they contain supplemental tables to accompany the results of statistical analyses reported in the main article.

**Measures of neuropsychological functioning**

All testing occurred in the morning in two 50-min counterbalanced sessions. Dunedin Study members completed a small number of neuropsychological measures at age 13, and then a larger battery at age 38.

Neuropsychological functioning at age 13 years.

We assessed *executive functioning* with the following test:

*Trails-B* is a test of scanning and tracking, divided attention, and mental flexibility, which involves drawing lines to connect consecutively numbered and lettered circles, alternating between numbers and letters. Scores represent the time, in seconds, to complete the test. Lower scores are better (Army Individual Test Battery, 1944). Test-retest reliability=0.73-0.89.

We assessed *memory* with the following tests:

*Rey Auditory Verbal Learning Test: Total Recall* is a test of verbal learning and memory that involves a five-trial presentation of a 15-word list and a one- time presentation of an interference list. Four trials of the 15-word list were administered due to time constraints. The Total Recall indexes the total number of words (0–60) recalled over four trials (the sum of words recalled across trials 1–4). Higher scores are better (Lezak, 2004); Test-retest reliability=0.86.

*Rey Auditory Verbal Learning Test: Delayed Recall* indexes the total number of words from the original list (0–15) recalled after a 25–30 minute delay. Higher scores are better (Lezak, 2004). Test-retest reliability=0.80.

Neuropsychological functioning at age 38 years.

We assessed *executive functioning* with the following tests:

*CANTAB Rapid Visual Information Processing: A‘ (A-prime)*, is a signal-detection measure that taps sustained attention, often called attentional vigilance. The participant scans for a 3-digit target sequence in a digit stream that is ongoing for 7 minutes, and responds whenever a target sequence is spotted. At the most difficult level, the participant scans simultaneously for two target sequences. Higher scores are better (Sahakian & Owen, 1992); Test-retest reliability=0.76.

*CANTAB Rapid Visual Information Processing: Total False Alarms* records impulsive jumping to respond too soon before the correct target digit sequence is complete. Because relatively few participants made numerous false alarms, this measure is categorical, coded 0=none, 1=1 false alarm, 2=2 or more false alarms. Lower scores are better (Sahakian & Owen, 1992); Test-retest reliability not available.

*WAIS Working Memory Index* was derived from the Arithmetic and Digit-span subtests. The Arithmetic subtest of WAIS requires working memory processes to be applied to orally presented verbal information. It involves numerical knowledge, short-term memory, attention, and concentration. Arithmetic problems are presented in story format. Performance requires holding information in short-term memory, accessing long-term memory to retrieve numerical rules of mathematical operation, and using the rules to manipulate the stored data. Items are arranged according to the level of difficulty and have time limits. The Digit-span subtest of WAIS is a test of memory span, attention/concentration, and ability to mentally manipulate information. The test requires listening to a sequence of digits read aloud and repeating them in forward, backward, and ascending order. Digit sequences range in length from 2 to 9 digits and are presented in order of difficulty. Higher scores on the WAIS Working Memory Index are better (Wechsler, 2008); Test-retest Reliability=0.90

*Wechsler Memory Scale-III: Months of the Year Backwards Test* is a test of attention and tracking. It requires reciting the months of the year in backwards order, starting with December. Responses were scored according to the instructions in the WMS-III manual. Scores ranged from 1 (poor performance) to 5 (good performance) and reflect both accuracy and speed. Higher scores are better (Wechsler, 1997); Test-retest Reliability=0.80

*Trails-B* is a test of scanning and tracking, divided attention, and mental flexibility, which involves drawing lines to connect consecutively numbered and lettered circles, alternating between numbers and letters. Scores represent the time, in seconds, to complete the test. Lower scores are better (Army Individual Test Battery, 1944). Test-retest reliability=0.73-0.89.

We assessed *memory* with the following tests:

*CANTAB Paired Associates Learning: First Trial Memory Score* is a test of visual memory and new learning. Boxes are displayed on the screen and are opened in a random order. One or more of them will contain a pattern. The patterns are then displayed in the middle of the screen, one at a time, and the subject must touch the box where the pattern was originally located. If the subject makes an error, the patterns are re-presented to remind the subject of their locations. The difficulty level increases through the test. The number of patterns increases across eight stages (i.e., two 1-pattern stages, two 2-pattern stages, two 3-pattern stages, one 6-pattern stage, one 8-pattern stage), which challenges even very able subjects. For each stage, up to 10 trials are presented until all the patterns are located correctly. For the First Trial Memory Score, the number of patterns correctly located after the first trial of each stage is summed across the stages completed (range 0–26, with 26 meaning that all the patterns were correctly located for all stages the first time). Higher scores are better (Sahakian & Owen, 1992); Test-retest reliability=0.87.

*CANTAB Paired Associates Learning: Total Errors* is based on the same protocol as above but considers the total number of errors (with an adjustment for each stage not attempted due to previous failure). Lower scores are better (Sahakian & Owen, 1992); Test-retest reliability=0.64.

*Wechsler Memory Scale-III: Verbal Paired Associates Total Recall* is a test of verbal learning and memory. Eight pairs of unrelated words (e.g., hat-sofa) are read aloud and followed by a recall task (one of the words from each word pair is given, and the associated word must be recalled). Four trials of the eight word-pairs are presented. Presentation of the word-pairs is randomized across trials. The total recall score represents the total number of words (0–32) recalled across four trials. Higher scores are better (Wechsler, 1997); Test-retest reliability=0.75.

*Wechsler Memory Scale-III: Verbal Paired Associates Delayed Recall* is based on the same protocol as above but represents the total number of words (0–8) recalled after a 25–35 minute delay. Higher scores are better (Wechsler, 1997); Test-retest reliability=0.73.

*Rey Auditory Verbal Learning Test: Total Recall* is a test of verbal learning and memory that involves a five-trial presentation of a 15-word list and a one- time presentation of an interference list. Four trials of the 15-word list were administered due to time constraints. The Total Recall indexes the total number of words (0–60) recalled over four trials (the sum of words recalled across trials 1–4). Higher scores are better (Lezak, 2004); Test-retest reliability=0.86.

*Rey Auditory Verbal Learning Test: Delayed Recall* indexes the total number of words from the original list (0–15) recalled after a 25–30 minute delay. Higher scores are better (Lezak, 2004). Test-retest reliability=0.80.

**Table S.1. Testing the “cognitive reserve hypothesis”: Neuropsychological test performance of Study members who did and did not receive a diagnosis of MDD by age 38.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Measures (All Administered at Age 13) | Never Depressed |  | Ever Depressed |   | Poisson Regression Results |
|  |  |
| *N* | Mean (95% CI) |  | *N* | Mean (95% CI) |   | IRR (95% CI) | p |
| *Executive Functioning* |  |  |  |  |  |  |  |  |
|  Trails B (time in seconds) | 368 | 35.5 (33.6, 37.3) |  | 286 | 37.5 (35.5, 39.4) |  | 1.00 (1.00, 1.01) | 0.053 |
| *Memory* |  |  |  |  |  |  |  |  |
|  Rey Auditory Verbal Learning Test -- Total words recalled | 375 | 40.6 (39.9, 41.3) |  | 290 | 41.3 (40.5, 42.1) |  | 0.99 (0.96, 1.02) | 0.551 |
|  Rey Auditory Verbal Learning Test -- Delayed recall | 374 | 9.9 (9.7, 10.2) |  | 289 | 10.0 (9.7, 10.4) |   | 1.00 (0.99, 1.01) | 0.733 |

*Notes.* This table includes Study members from the full analytic sample (see **Figure 1A**) who completed neuropsychological testing at age 13. To ensure that cognitive functioning at age 13 was being used to predict *future* MDD, Study members who received a diagnosis of MDD prior to age 18 were excluded from this analysis. The “Never Depressed” group represents Study members who did not ever meet criteria for MDD between the ages of 18 and 38. The “Ever Depressed” group represents Study members who did meet diagnostic criteria for MDD during this same period. Relative risk (IRR) of MDD was calculated using a modified Poisson regression model with robust standard errors (Zou, 2004). Significant group differences (p < 0.05) are shown in **bold**.

**Table S.2. Testing the “scarring hypothesis”: Neuropsychological test performance of Study members who were not diagnosed with past-year MDD at age 38, by lifetime diagnostic history.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Measures (All Administered at Age 38) | Never Depressed |  | Ever Depressed |   |   |
|  |  |
| *N* | Mean (95% CI) |  | *N* | Mean (95% CI) |   | *F* | p |
| *Executive Functioning* |  |  |  |  |  |  |  |  |
|  Rapid Visual Processing -- A prime | 271 | 0.91 (0.91, 0.92) |  | 458 | 0.91 (0.91, 0.92) |  | 0.01 | 0.916 |
|  Rapid Visual Processing -- False alarms | 271 | 0.23 (0.18, 0.30) |  | 460 | 0.25 (0.20, 0.30) |  | 0.11 | 0.743 |
|  WAIS-IV Working Memory Index  | 277 | 103.5 (101.6, 105.4) |  | 467 | 104.4 (103.0, 105.7) |  | 0.03 | 0.867 |
|  WMS-III Months of the Year Backwards | 276 | 3.12 (2.96, 3.28) |  | 467 | 3.12 (3.00, 3.25) |  | 1.93 | 0.165 |
|  Trails B (time in seconds) | 277 | 63.9 (61.5, 66.3) |  | 466 | 63.4 (61.6, 65.2) |  | 0.81 | 0.367 |
| *Memory* |  |  |  |  |  |  |  |  |
|  CANTAB Paired Associates -- First Trial | 273 | 20.2 (19.8, 20.7) |  | 460 | 20.3 (20.0, 20.6) |  | 0.21 | 0.649 |
|  CANTAB Paired Associates -- Total Errors | 273 | 12.3 (9.8, 14.8) |  | 461 | 11.0 (9.8, 12.1) |  | 1.83 | 0.176 |
|  Verbal Paired Associates -- Total N correct | 276 | 17.1 (16.1, 18.1) |  | 467 | 15.6 (14.9, 16.4) |  | 3.07 | 0.080 |
|  **Verbal Paired Associates -- Delayed recall total** | **276** | **5.6 (5.3, 5.9)** |  | **465** | **5.1 (4.8, 5.3)** |  | **4.40** | **0.036** |
|  Rey Auditory Verbal Learning Test -- Total words recalled | 275 | 38.7 (37.8, 39.6) |  | 467 | 37.1 (36.4, 37.8) |  | 1.32 | 0.252 |
|  Rey Auditory Verbal Learning Test -- Delayed recall | 276 | 9.4 (9.1, 9.8) |  | 467 | 9.0 (8.7, 9.3) |   | 0.11 | 0.745 |

*Notes.* This table includes only those Study members who (a) completed neuropsychological testing at age 38, (b) were not diagnosed with past-year MDD at age 38 and (b) had present data for adult IQ. The “no history of MDD” group represents Study members who had never met criteria for MDD. The “past history of MDD” group represents Study members who had met diagnostic criteria for MDD during a previous wave but no longer met criteria at age 38. Means were compared across diagnostic groups through a series of one-way ANOVAs, controlling for sex. Significant group differences (p < 0.05) are shown in **bold**.

**References**

Army Individual Test Battery. (1944). *Manual and directions for scoring*. Washington, DC: War Department.

Lezak, M. D. (2004). *Neuropsychological Assessment* (4th ed.). New York, NY: Oxford University Press.

Sahakian, B. J., & Owen, A. M. (1992). Computerized assessment in neuropsychiatry using CANTAB: discussion paper. *Journal of the Royal Society of Medicine*, *85*, 399–402.

Wechsler, D. (1997). *Wechsler memory scale* (3rd ed.). San Antonio, TX: Psychological Corporation.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale (4th ed.)*. San Antonio, TX: Pearson Assessment.

Zou, G. (2004). A modified Poisson regression approach to prospective studies with binary data. *American Journal of Epidemiology*, *159*(7), 702–706. https://doi.org/10.1093/aje/kwh090