Add Health
The National Longitudinal Study of Adolescent to Adult Health

Report prepared by

Allison E. Aiello

Rebecca C. Stebbins

Cherese N.B. Parker

Farizah I. Rob

Robert A. Hummer

TestMyBrain User Guide

Add CAPS
Add Health Cognitive Assesment, Physical, and Sensory Function

UNC CAROLINA POPULATION CENTER

CAROLINA POPULATION CENTER | CAROLINA SQUARE - SUITE 210 | 123 WEST FRANKLIN STREET | CHAPEL HILL, NC 27516

https://doi.org/10.17615/axvk-sr06

Acknowledgement:

Data collection and dissemination for Wave VI of Add Health was funded by cooperative agreements U01 AG071448 (PI: Robert A. Hummer) and U01 AG071450 (MPIs: Robert A. Hummer and Allison E. Aiello) from the National Institute on Aging to the University of North Carolina at Chapel Hill, with cooperative funding from five other institutes and offices at the National Institutes of Health: the Eunice Kennedy Shriver National Institute of Child Health and Human Development, the National Institute on Minority Health and Health Disparities, the National Institute on Drug Abuse, the Office of Disease Prevention, and the Office of Behavioral and Social Science Research. Waves I-V of Add Health were funded by grant P01 HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

Important Note: The above acknowledgement should be included in all presentations and publications using data from Wave VI of Add Health.

Citation for User Guide:

Aiello A.E., Stebbins R.C., Parker C.N.B., Rob F.I., and Hummer R.A. 2025. *TestMyBrain User Guide.* Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill. Available from: https://doi.org/10.17615/axvk-sr06

Table of Contents

1. Introduction

The National Longitudinal Study of Adolescent to Adult Health (Add Health) is a nationally representative sample of U.S. adolescents who were in grades 7-12 during the 1994-1995 school year. Using a complex, school-based cluster-sampling frame, researchers selected high school and feeder school pairs from 80 communities across the United States and drew a sex- and grade-stratified random sample of 20,745 adolescents for inclusion in the study. This sample has been followed from adolescence into early midlife across six waves of data collection to date, with the most recent wave of data collection (Wave VI) taking place between 2022 and 2025 when participants were ages 39 to 51, with an average age of 44.

Over the years, Add Health has collected a wealth of information from participants and their parents about demographic characteristics, familial structures, social relationships, health behaviors, cognition, physical and mental health status, medication usage, and health care access. Add Health also has collected anthropometric, cardiovascular, metabolic, renal, hepatic, inflammatory/immune, infectious, neurodegenerative, and multi-omic biomarkers from participants. In addition, Add Health has merged multilevel contextual data about the economic, school, neighborhood, policy, and environmental contexts in which the participants are embedded to the core survey and biological data at each wave. The Add Health project thereby provides researchers with rich opportunities to explore the causes and consequences of health status across multiple contextual domains as individuals age across the life course.

Given the breadth of measures and longitudinal design, Add Health offers a unique opportunity to explore the early origins of cognitive functioning and to track changes in dementia risk within a nationally representative cohort that has been followed since early adolescence. A key objective of Wave VI of Add Health was to significantly expand the measurement of cognitive domains. This wave introduced the Add Health Cognitive Assessment, Physical, and Sensory Function (Add CAPS) battery, designed to support long-term tracking of cognitive, physical, and sensory function. These data provide a foundation for identifying early midlife signs of cognitive impairment that may help map future risk for Alzheimer's disease and Alzheimer's disease related dementias (AD/ADRD).

The inclusion of *TestMyBrain* (TMB) reflects our commitment to collecting web-based cognitive measures in early midlife among a nationally representative sample. This approach also enables comparison with standard in-person assessments, helping to evaluate the feasibility and value of remote cognitive testing at scale in Add Health.

2. General Overview of TestMyBrain

TMB is an automated, self-administered, web-based, digital cognitive assessment platform run by the Many Brains Project (a 501(c)(3) non-profit; please see https://www.manybrains.net). TMB was selected to achieve the Add Health team's goal of integrating a 15-minute digital cognition

battery into the Add Health Wave VI survey. The Add Health Wave VI custom TMB cognition battery was designed by selecting from existing TMB cognitive assessments that focused on the following cognitive domains: Verbal Episodic Memory, Cognitive Processing Speed, Executive Function, and Working Memory. Through a comprehensive literature review, evaluation of available digital cognitive assessments, and consultation with neuropsychologists, these were identified as key domains of cognition for the 39–51 years-old Add Health Wave VI participants.

Each TMB testing battery consisted of a learning phase and 4 exercises, in the following order:

- Learning Phase: TMB Verbal Paired Associates - Word pairs
- TMB Digit Symbol Matching Test
- Testing Phase: TMB Verbal Paired Associates - Recall
- TMB Gradual Onset Continuous Performance Test
- TMB Digit Span - Backwards

Add Health Wave VI was divided into two nationally representative subsamples: Sample 1 and Sample 2 (as detailed in the Sampling and Mixed-Mode Survey Design User Guide; see Hummer et al. 2025). Both Sample 1 and Sample 2 received the same TMB cognitive assessments.

The TMB battery was integrated into the same section of the Add Health Wave VI survey for both subsamples. However, administration differed by mode of participation. Sample 1 participants completed the assessments independently through the web-based platform, embedded directly within the survey and accessible via any internet-connected device. In contrast, Sample 2 participants completed the assessments during their in-person interviews using a custom-built progressive web application, run offline on standard field interviewer laptops. While the TMB exercises were self-administered, all in-person interviews were facilitated by a trained field interviewer working directly with the Wave VI participant.

A small portion of Sample 1 participants (N=366) who did not respond to requests to complete the web survey subsequently completed an in-person interview instead. Similar to Sample 2, these participants completed TMB using the progressive web application during the in-person interview. In contrast, Sample 2 participants who were unable to complete the progressive web application during the in-person interview due to an administration error in the field were emailed a link to complete TMB via the web-based platform after the in-person interview was complete (N=11). Additionally, 70 Sample 2 participants who we were unable to interview in-person, completed the survey online. All participants have a variable indicating which TMB administration mode (online vs. progressive web application) was used: CB1MODE.

Differences in device type (desktop, laptop, tablet, or smartphone) and mode of data input (keyboard, mouse, touch-screen) may impact test results. Therefore, the variable PLATFORM, which reports device type, is provided in the wave6.sas7bdat data, while variables indicating input type are provided in accb1_6.sas7bdat: C6VPA0, C6DSM0, C6GCPT0, and C6BDS0.

3. TMB Verbal Paired Associates - Recall

3.1 Rationale

TMB Verbal Paired Associates - Recall is an automated self-administered digital cognitive assessment that measures verbal episodic memory. The domain of verbal episodic memory is important for the Add Health Wave VI study population because it is one of the first cognitive functions impacted by aging (Corte et al, 2019). The TMB Verbal Paired Associates – Recall serves as a baseline measurement for how well Add Health Wave VI participants can learn and remember associations between word pairs.

During the learning phase, participants were shown the 25 word-pairs, one pair at a time. During the Testing Phase, participants were shown one word from each pair and must choose which word, from the provided list of options, was paired with the displayed word. Further details on psychometric properties and of the Verbal Paired Associates are available in Appendix A.

3.2 Measurement and Protocol

The Verbal Paired Associates – Recall task consisted of two phases with a short delay between each phase:

- Learning Phase
  - Learning – Participants are visually presented with a series of word pairs on the screen (for example, "SKY-BLUE"). See Figure 1. Each pair is shown one at a time for a standard time interval. A message on the screen instructs participants to remember which words go together because they will be tested later on their memory for these pairings.

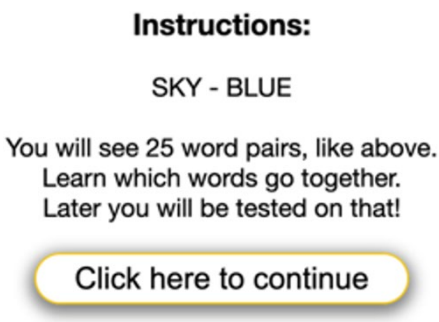


Figure 1. Screenshot of TMB Verbal Paired Associates Instruction Screen

- Delay and Distractor – After all pairs have been shown, there is a brief delay of about 2 minutes before the memory test. During this delay, the participant engages in an unrelated short task to prevent rehearsal or interference.

- Testing Phase – In the test phase, participants are shown one word from each pair and must choose the correct mate from a set of four options. Specifically, on each trial a previously learned cue word is displayed (e.g., "OFFICER – ?") with four possible answer words listed. The participant taps or clicks on the word they believe was paired with the cue.
    - During the assessment, no feedback is given about performance.

3.3 Variable Construction

The following set of variables for Verbal Paired Associates were considered to be relevant for Add Health users and have been included in the dataset.

3.3.1. Score (C6VPA1)

The number of correct responses in 'test' trials.

3.3.2. Accuracy (C6VPA2)

The proportion of correct responses for 'test' trials. **This is the recommended primary research outcome.**

3.3.3. Duration (C6VPA3)

The duration it took for the participant to start and finish the Verbal Paired Associates task (not including the pre-exercise component), in seconds.

3.3.4. QC Flag – Verbal Paired Associates (C6VPAQ4)

A flag variable was created to indicate cases with implausibly short reaction times (median reaction time < 1000 ms).

3.4 Quality Control

Throughout Wave VI, weekly reports were generated for individual TMB assessments, incorporating qualitative flags to capture occurrences such as technical issues, distractions/interruptions, and other issues during administration. For the Verbal Paired Associates - Recall task, key measures such as duration, reaction time, accuracy, and score were statistically summarized both overall and across demographics to identify outliers and implausible values. Reaction times deemed implausibly fast (median reaction time < 1000 ms) were flagged weekly, and occurrences remained rare (Sample 1 = 1.1% and Sample 2 = 1.0%).

3.5 Recommended variable usage

The recommended primary research outcome for Verbal Paired Associates – Recall is **accuracy**. A *higher* value of the **accuracy** variable indicates better cognitive performance.

4. TMB Digit Symbol Matching Test

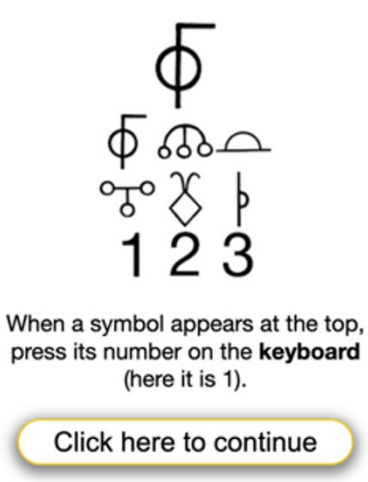
4.1 Rationale

TMB Digit Symbol Matching Test is an automated self-administered digital cognitive assessment that measures processing speed. The domain of processing speed, or the amount

of time required to complete a task, is important for the Add Health Wave VI study population because it is sensitive to impairments in neurological function and follows a well-defined trajectory over the lifespan (Carlozzi et al, 2015). Processing speed increases throughout childhood and peaks in early adulthood, declining throughout midlife and older age.

4.2 Measurement and Protocol

In the Digit Symbol Matching Test, the participant sees a key or legend at the top of the screen that pairs a set of symbols with digits. Specifically, six abstract symbols are each associated with a digit 1 through 3 and this key remains visible throughout the task. During the test, individual symbols are presented one by one in a target area, and the participant must quickly select the corresponding number for that symbol based on the reference key.



The participant is first shown an instruction screen explaining the symbol-digit mappings, as shown in Figure 2. Participants then get a brief practice to ensure they understand how to use the key before the timed test begins.

The timed test then begins. A symbol appears, and the participant selects the number corresponding to that symbol. Upon responding, the next symbol immediately appears. The goal is to match as many symbols to their digits as possible within 90 seconds. Each symbol remains on screen until the participant responds.

Figure 2. Screenshot from TMB Digit Symbol Matching Test instruction screen.

The TMB system records each response and whether it was correct. Scoring is based on the number of correct matches completed in the time limit.

4.3 Variable Construction

The following set of summary-level variables for Digit Symbol Matching were considered to be relevant for Add Health users and have been included in the dataset. Further details on psychometric properties and validation results of the Digit Symbol Matching Test are available in Appendix B.

4.3.1. Num_correct (C6DSM1)

The number of items the participant got correct.

4.3.2. Num_responses (C6DSM2)

The total number of items presented to the participant.

4.3.3. Score (C6DSM3)

The number of items the participant matched correctly.

4.3.4. Median Reaction Time (C6DSM4)

The median reaction time of correct responses during the assessment. Lower values (faster median reaction time) represent better performance. **This is the recommended primary research outcome.**

4.3.5. Duration (C6DSM5)

The duration it took for the participant to start and finish the Digit Symbol Matching test, in milliseconds.

4.3.6. QC Flag – Digit Symbol Matching Test (C6DSMQ6)

Scores suggesting lack of engagement, careless responding, guessing, or implausibly fast reaction times were flagged. Digit Symbol Matching test data was flagged if at least one of the following criteria was met:

1. Number of correct items < 6
2. Proportion of correct items (num_correct/num_responses) < 0.5
3. Median reaction time < 300 milliseconds
4. Standard deviation of reaction times of correct responses in 'test' trials/Median reaction times of correct responses to 'test' trials > 2

4.4 Quality Control

Throughout Wave VI, weekly reports were generated for individual TMB assessments, incorporating qualitative flags to capture occurrences such as technical issues, distractions/interruptions, and other issues during administration. For the Digit Symbol Matching Test, key measures such as median reaction time, duration of the test, and the number of correct responses were statistically summarized both overall and across demographics to identify outliers and implausible values. Digit Symbol Matching Test data that was considered indicative of participants not being engaged with the task, responding carelessly or guessing, and implausibly fast reaction times were flagged weekly (Sample 1 = 3.3% and Sample 2 = 3.4%, at the end of data collection).

4.5 Recommended variable usage

The recommended primary research outcome for the Digit Symbol Matching Test is **median reaction time [C6DSM4]**. A *lower* value of the **median reaction time** variable indicates better cognitive performance.

5. TMB Gradual Onset Continuous Performance Test

5.1 Rationale

The TMB Gradual Onset Continuous Performance Test is an automated self-administered digital cognitive assessment that measures executive function in the form of cognitive control, sustained attention, and response inhibition. The domain of executive function is important to measure at this stage in the life course for the Add Health Wave VI study population, because executive dysfunction impacts activities of daily living in aging populations, which reduces autonomy and quality of life (Faria et al, 2015).

- This is a Go/No-Go test (a task that tests both quick responses and the ability to stop yourself when needed) of cognitive control, sustained attention, and response inhibition.
- This test manipulates stimulus presentation in such a way that it rapidly exhausts attentional resources, making it very sensitive to individual differences in vigilance while also being very brief, allowing reliable measures in 3-4 minutes as opposed to 15-20 minutes.

5.2 Measurement and Protocol

The Gradual Onset Continuous Performance Test (GCPT) uses gradually transitioning images to require constant vigilance. In the Add Health Wave VI implementation via TMB, participants see a rapid sequence of 300 photographic images that continuously morph from one to the next. The two categories of images are city scenes (89.3% of trials) and mountain scenes (10.7% of trials). See Figure 3 for an example image for each category. The participants are instructed to click whenever a city image appears and to do nothing when a mountain image appears.

Figure 3. Screenshot of TMB GCPT instruction screen.

Key details of the protocol:

- Images of cities and mountains are blended into one another in a continuous stream. Each image gradually fades into the next over a short period. At any given moment, an image can become apparent to the participant.
- Before the test starts, participants are given on-screen instructions and a short practice. They learn to click for city and not for mountain.
- The participant must maintain focus and respond quickly to each city scene. If a mountain scene appears, they must withhold the response (i.e., not click).
- During the test, no feedback is given about errors.

5.3 Variable Construction

The following set of variables for GCPT were considered to be relevant for Add Health users and have been included in the dataset. Additional details on psychometric properties and validation techniques for designing the test have been documented in Appendix C by the Many Brains team.

5.3.1 Accuracy (C6GCPT1)

The proportion of correct responses to 'test' trials.

5.3.2 Median reaction time (C6GCPT2)

The median reaction time of correct responses to 'test' trials, in milliseconds.

5.3.3 CEprop (C6GCPT3)

The "commission error" proportion. This measures when a participant clicks when they should not.

5.3.4 OEprop (C6GCPT4)

The "omission error" proportion. This measures when a participant does not click when they should.

5.3.5 dPrime (C6GCPT5)

This is a measure of response accuracy unaffected by response bias (user's sensitivity), which represents the participant's ability to withhold responses to mountains while making responses to the more prevalent city images. Higher values indicate better performance. **This is the recommended primary research outcome.** It is computed using the formula:

pHit = 1 - CEprop (subtract 0.5 error if CEprop = 0)

pFA = OEprop (add 0.5 error if OE = 0)

dPrime = norminv(pHit) - norminv(pFA)

5.3.6 Response Bias (C6GCPT6)

This is a measure of the participant's bias. A larger response bias value indicates a greater tendency to press a key regardless of the picture type, suggesting impulsivity, and is computed as follows:

crit = (-1 * (norminv(pHit) + norminv(pFA)) / 2

5.3.7 Score (C6GCPT7)

Participant's returned performance score, computed as follows: (1 - CEprop)*100

5.3.8 Duration (C6GCPT8)

The duration it took the participant to start and finish the Gradual Onset Continuous Performance test, in seconds.

5.3.9 QC Flag - GCPT (C6GCPTQ9)

Data indicative of participants not being engaged with the task were flagged, based on the omission error proportion. If OEprop >= 0.5, representing that the participant did not click when they should have more than half the time, the score was flagged.

5.4 Quality Control

Throughout Wave VI, weekly reports were generated for individual TMB assessments, incorporating qualitative flags to capture occurrences such as technical issues, distractions/interruptions, and other issues during administration. The accuracy in GCPT was statistically summarized both overall and across demographics to identify outliers and implausible values. GCPT data that was considered indicative of participants not being engaged with the task were flagged weekly (Sample 1 = 2.0% and Sample 2 = 2.7%, at the end of data collection).

5.5 Recommended variable usage

The recommended primary research outcome for Gradual Onset Continuous Performance Test is **dprime**. A *higher* value of the **dprime** variable indicates better cognitive performance.

6. TMB Backward Digit Span

6.1 Rationale

TMB Backwards Digit Span Test is an automated self-administered digital cognitive assessment that measures working memory by assessing a participant's ability to remember a sequence of numbers in reverse order. The domain of working memory is important for the Add Health Wave VI study population because it is commonly used to monitor cognitive health and detect early signs of cognitive impairment (Kirova et al, 2015). This test was adapted from the Wechsler Adult Intelligence Scales.

6.2 Measurement and Protocol

The TMB version of the Backward Digit Span task was digitally administered to Add Health Wave VI participants as part of the TMB cognitive battery. The digitally administered backward digit span closely mirrors the traditional procedure but uses visual presentation and keyboard/touchscreen input. The sequence of administration was as follows:

- Participants first received on-screen instructions explaining that they would see numbers and needed to remember them *in reverse order*. An example was provided (for instance, if shown "3-7-2", they should input "2-7-3"). They could practice a

couple of simple sequences to ensure they understood the concept of “backward” recall.

- Digits were presented one at a time in the center of the screen, each for about 1 second. See Figure 4. A predetermined sequence of digits, ranging from 1–9, was displayed. After the sequence was shown, there was a prompt such as “Type the numbers in reverse order now:” and an input field. Note: when a touchscreen-capable device was detected, the user was presented with the following screener at the onset of the test: Please choose how you will respond in this test: 1) to use a keyboard, press any key now. 2) to use a touchscreen, touch the screen now.
- Participant Response: The participant then typed the digits in the reverse order from what they saw. For example, if the screen showed “5-8” sequentially, the participant would enter “8-5”.
- Two trials were provided for each sequence-length of numbers, and the length increased by one digit on the next trial if the participant got at least one of those two trials correct. This continues to 4-digit, 5-digit sequences, etc., each length with two chances, up to 11-digits.
- If the participant failed both trials at a given length, the task ended. The longest span length at which they correctly recalled a sequence backwards is the score.

The maximum span tested in the TMB version is higher than the traditional in-person version, as the TMB protocol allows up to 11 digits.

The TMB system automatically calculates the Backward Digit Span score, which is the length of the longest list recalled (i.e. the number of digits recalled) correctly. For example, if a participant correctly reversed one 6-digit sequence but failed both 7-digit sequences, their score would be 6.

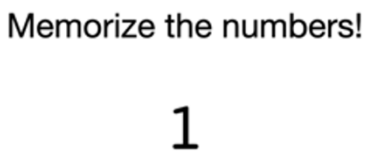


Figure 1. Screenshot of TMB Backward Digit Span screen.

6.3 Variable Construction

The following set of variables for the Backward Digit Span are available for Add Health users. The data dictionary of all variables created by TMB for the Backward Digit Span test can be found in Appendix D.

6.3.1. NumCorrect (C6BDS1)

The number of items the participant got correct.

6.3.2. Span (C6BDS2)

The maximum length of numbers the participant got correct. **This is the recommended primary research outcome.**

6.3.3. Duration (C6BDS3)

The duration it took for the participant to start and finish the Backward Digit Span test.

6.3.4. QC Flag – Backward Digit Span (C6BDSQ4)

If the maximum length of numbers the participant got correct was less than 2 (span < 2), then the data was flagged as an indicator of participant's inability to complete task or follow directions for the Backward Digit Span test.

6.4 Quality Control

Throughout Wave VI, weekly reports were generated for individual TMB assessments, incorporating qualitative flags to capture occurrences such as technical issues, distractions/interruptions, and other issues during administration. The completion rate for Backward Digit Span is lowest compared to the previous three tests of the TMB cognitive battery due to a recurring technical issue with the practice round. Score and total duration of the Backward Digit Span test were analyzed weekly, both overall and across demographics to identify outliers and implausible values. To assess validity, comparisons were made between the interviewer-administered Backward Digit Span (see the Interviewer-Administered Word Recall and Backward Digit Span User Guide) and the TMB-Backward Digit Span (r = 0.53). Backward Digit Span data indicative of participant's inability to complete task or follow directions were flagged weekly (Sample 1 = 2.1%, and Sample 2 = 4.3%, at the end of data collection). Additionally, we flagged cases where the number of practice trials > 10 (Sample 2 = 3.9%).

6.5 Recommended variable usage

The recommended primary research outcome for Backward Digit Span is **span-** which represents the length of the longest backward sequence of numbers that were remembered correctly. A *higher* value of the **span** variable indicates better cognitive performance.

7. Additional Quality Control Flags

7.1 Qualitative Flag 1 (C6CB1F1)

Tests were flagged with Qualitative Flag 1 if a distraction or interruption occurred during the test.

7.2 Qualitative Flag 2 (C6CB1F2)

Tests were flagged with Qualitative Flag 2 when an FI capture issue was reported.

7.3 Qualitative Flag 3 (C6CB1F3)

Tests were flagged with Qualitative Flag 3 when a technical issue occurred.

7.4 Qualitative Flag 4 (C6CB1F4)

Tests were flagged with Qualitative Flag 4 when the participant reported physical impairment or pain.

7.5 Qualitative Flag 5 (C6CB1F5)

Tests were flagged with Qualitative Flag 5 when an administration or instruction issue occurred.

7.6 Qualitative Flag 6 (C6CB1F6)

Tests were flagged with Qualitative Flag 6 when the participant reported fatigue or displayed low effort.

8. Survey Memory Questions and Other Cognitive Data

Of note, subjective reports of memory were asked of both Sample 1 and Sample 2 participants as part of the main Add Health Wave VI survey. These data can be found in Section 4: Health Care and Illness dataset, variables H6DA18B, H6DA18C, H6DA18D, H6DA18E, and H6DA18F. Additional cognitive function measures, completed by Sample 2 participants only, can be found in our additional Add CAPS Users Guides: NIH Toolbox Cognition User Guide, Animal Fluency User Guide, and Interviewer-Administered Word Recall and Backward Digit Span User Guide.

9. The TestMyBrain Data File (accb1_6.sas7bdat)

9.1 Structure

The structure of the disseminated TMB data file is flat. This means that it is a participant-level data file, wherein each participant has one and only one record. The participant identifier (AID) will appear in the data file only once. For every participant, SAMPLE indicates which sample a participant was part of, and CB1MODE indicates with which mode of TMB the participant completed the tests.

9.2 Contents

The TMB data file includes the variables below, which are described in the corresponding codebook documentation that also contains frequencies.

Variable Name	**Variable Description**
AID	Participant Identifier
SAMPLE	Sample Assignment-W6
CB1MODE	Cognitive Battery 1 Mode-W6
C6VPA0	Verbal Paired Associates mode of input
C6VPA1	Verbal Paired Associates score
C6VPA2	Verbal Paired Associates accuracy
C6VPA3	Verbal Paired Associates duration
C6VPAQ4	Verbal Paired Associates implausibly fast reaction time
C6DSM0	Digit Symbol Match mode of input
C6DSM1	Digit Symbol Match number correct
C6DSM2	Digit Symbol Match number of responses
C6DSM3	Digit Symbol Match score
C6DSM4	Digit Symbol Match median reaction time
C6DSM5	Digit Symbol Match duration
C6DSMQ6	Digit Symbol Match not engaged/careless/guessing/fast
C6GCPT0	GCPT mode of input

C6GCPT1	GCPT accuracy
C6GCPT2	GCPT median reaction time
C6GCPT3	GCPT CEprop
C6GCPT4	GCPT OEprop
C6GCPT5	GCPT dprime
C6GCPT6	GCPT participant bias
C6GCPT7	GCPT score
C6GCPT8	GCPT duration
C6GCPTQ9	GCPT not engaged
C6BDS0	Backward Digit mode of input
C6BDS1	Backward Digit number correct
C6BDS2	Backward Digit span
C6BDS3	Backward Digit duration
C6BDSQ4	Backward Digit unable to complete/follow directions
C6CB1F1	distraction or interruption
C6CB1F2	field interviewer capture issues
C6CB1F3	technical issue
C6CB1F4	physical impairment or pain
C6CB1F5	administration or instruction
C6CB1F6	fatigue or low effort
C6CB1TS	Cognitive battery 1 duration, seconds
C6CB1TF	Was the test duration as expected, shorter than expected, or longer than expected
H6CB1T1	technical difficulties
H6CB1T2	participant took a break
H6CB1T3	interruption
H6CB1T4	did not complete test
H6CB1T5	did not understand instructions
H6CB1T6	physical or cognitive disability
H6CB1T7	other
H6CB1T8	refused to participate

10.References

Faria CA, Alves HVD, Charchat-Fichman H. The most frequently used tests for assessing executive functions in aging. Dement Neuropsychol. 2015;9(2):149-155. doi:10.1590/1980-57642015DN92000009

Hummer RA, Dean SC, Liao D, Cooney D, Jackson C, Griffiths A, Aiello AE. Wave VI Sampling and Mixed-Mode Survey Design. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill; 2025.

Kirova AM, Bays RB, Lagalwar S. Working memory and executive function decline across normal aging, mild cognitive impairment, and Alzheimer's disease. Biomed Res Int. 2015;2015:748212.

doi:10.1155/2015/748212

La Corte V, Sperduti M, Abichou K, Piolino P. Episodic memory assessment and remediation in normal and pathological aging using virtual reality: a mini review. Front Psychol. 2019;10:173. doi:10.3389/fpsyg.2019.00173

Carlozzi NE, Beaumont JL, Tulsky DS, Gershon RC. The NIH Toolbox Pattern Comparison Processing Speed Test: normative data. Arch Clin Neuropsychol. 2015;30(5):359-368. doi:10.1093/arclin/acv031

Appendix A

Cognitive Task Overview: TestMyBrain Verbal Paired Associates Memory

Contact: Info@ManyBrains.net
ManyBrains.net
TestMyBrain.org

TMB Test Name: TestMyBrain Verbal Paired Associates Memory
Test Demos: Standard Concrete Study | Test
Document Version: May.29.2024
Acknowledgments: Christopher Chabris, Geisinger Health

The Many Brains Project

The Many Brains Project, is a 501(c)(3) non-profit that supports TestMyBrain (TMB) in collaboration with the Laboratory for Brain and Cognitive Health Technology at McLean Hospital and Harvard Medical School. We currently support many different types of research studies through our infrastructure for cognitive assessment - these range in size from small lab-based pilot studies to large longitudinal, multisite clinical research studies with tens of thousands of participants. As TestMyBrain has been continuously in operation since 2008, we provide a stable and secure platform for hosting and delivering mobile and web-based cognitive assessment protocols. Through TestMyBrain.org, data have been collected from over 3.7 million participants in a *citizen science* framework that includes structured return of research results toward the development, validation, and normative characterization of cognitive measures. We currently support research and education at over 2,000 sites worldwide engaged in digital neuropsychological assessment.

CITATION

Please credit The Many Brains Project and TestMyBrain in any papers, posters, or publications related to the TMB tests or data collected by TMB tests.

- Example:
  - All tasks were selected from and hosted on The Many Brains Project's web-based cognitive testing platform, TestMyBrain (Germine et al., 2012; The Many Brains Project).
    - Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, *19*(5), 847-857.
    - The Many Brains Project. *TestMyBrain Cognitive Tests*. URL: www.manybrains.net

Test Overview

Background:
TMB Verbal Paired Associates Memory (Germine et al., 2012; Singh et al., 2021, Wilmer et al., 2012) is a verbal episodic memory task, adapted from the Wechsler Memory Scale-III (Wechsler, 1997) for remote, unsupervised administration.

Task Parameters:
In the standard-length versions of the test, participants are visually presented with 25 pairs of words (e.g., truck | lemon), and informed they will later be tested on which words were paired together. Word pairs are presented sequentially for 3000 ms each, with a 500 ms interstimulus interval between pairs. After a delay of approximately 1.5-2.5 minutes, during which another brief test is typically completed, participants are sequentially presented with one word from each of the studied word pairs, and asked to identify which word was previously paired with it by selecting the correct word from a list of four response options. On each trial, one of the wrong answers is a correct response to another trial, one of the wrong answers is an incorrect response to another trial, and one incorrect response appears only during the current trial. Participants completed one unscored practice trial before beginning the 25 test trials. Participants have 10 seconds to make a response to each trial, after which a message appears warning the participant to respond more quickly; following the timeout message, the trial is repeated.

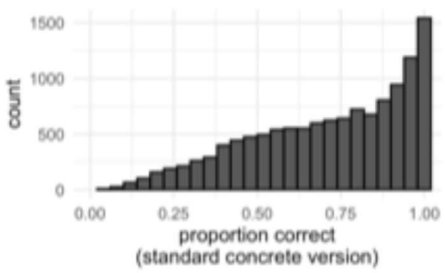
Primary Outcome:
The suggested primary outcome for the test is the proportion of test trials answered correctly (accuracy).

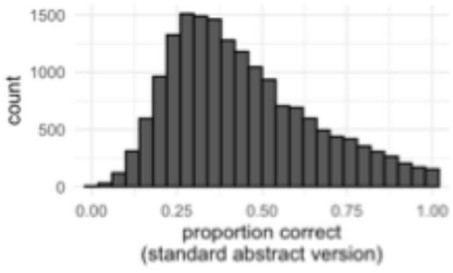
User Input:
Participants respond either by touching their selections (touch-compatible devices) or clicking their selections with a mouse press.

Alternate Task Versions: Multiple concrete and abstract versions of the test are available, at both standard and EMA-length.

Psychometrics:

- **Reliability**: In single-session testing, variation in accuracy between participants has a split-half reliability of .89 for the standard concrete version, and .83 for the standard abstract version.
- **Score distribution**:





References:

Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review, 19*(5), 847-857.

Singh, S., Strong, R. W., Jung, L., Li, F. H., Grinspoon, L., Scheuer, L. S., Passell, E. J., Martini, P., Chaytor, N., Soble, J. R., & Germine, L. (2021). The TestMyBrain Digital Neuropsychology Toolkit: Development and Psychometric Characteristics. *Journal of Clinical and Experimental Neuropsychology, 43*(8), 786-795.

Wechsler, D. (1997). Wechsler Memory Scale-III. San Antonio, TX: Psychological Corporation.

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Gerbasi, M., & Nakayama, K. (2012). Capturing specific abilities as a window into human individuality: The example of face recognition. *Cognitive Neuropsychology, 29*(5-6), 360-392.

Appendix B

Cognitive Task Overview: TestMyBrain Digit Symbol Matching

Contact: Info@ManyBrains.net
ManyBrains.net
TestMyBrain.org

TMB Test Name: TestMyBrain Digit Symbol Matching
Test Demo: standard version
Document Version: April.22.2024

The Many Brains Project

The Many Brains Project is a US-based 501(c)3 non-profit focused on the development of digital cognitive testing tools. We currently support many different types of research studies through our infrastructure for cognitive assessment - these range in size from small lab-based pilot studies to large longitudinal, multisite clinical research studies with tens of thousands of participants. As TestMyBrain.org has been continuously in operation since 2008, we provide a stable and secure platform for hosting and delivering mobile and web-based cognitive assessment protocols. Through TestMyBrain.org, data have been collected from over 2.5 million participants in a *citizen science* framework that includes structured return of research results toward the development, validation, and normative characterization of cognitive measures. We currently support research and education at over 400 sites worldwide as well as support for over 1200 clinicians or clinical sites engaged in remote digital neuropsychological assessment. For more information contact info@manybrains.net.

CITATION

Please credit The Many Brains Project and TestMyBrain in any papers, posters, or publications related to the TMB tests or data collected by TMB tests.

- Example:
  - All tasks were selected from and hosted on The Many Brains Project's web-based cognitive testing platform, TestMyBrain (Germine et al., 2012; The Many Brains Project).
    - Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, *19*(5), 847-857.
    - The Many Brains Project. *TestMyBrain Cognitive Tests*. URL: www.manybrains.net

Test Overview

Background:
TMB Digit Symbol Matching (Chaytor et al., 2021; D'Ardenne et al., 2020; Hartshorne & Germine, 2015; Hawks et al., 2023; Pozo et al., 2022; Singh et al., 2021; Singh et al., 2023) is a processing speed test adapted from the WAIS-III (Wechsler, 1997) for remote, digital administration.

Task Parameters:
Participants are presented with six symbols, each of which is paired with a single digit between 1-3 (i.e., two symbols are paired with each digit). These digit-symbol pairings remain visible throughout the duration of the test. Individual probe symbols are sequentially presented above the digit-symbol pairings, to which participants respond by selecting the corresponding digit as quickly as possible for 90 seconds; each probe symbol remains visible until participants make a response, or until 90 seconds have elapsed from the beginning of the test trials. Participants complete three practice probes before beginning the 90 seconds of test probes.

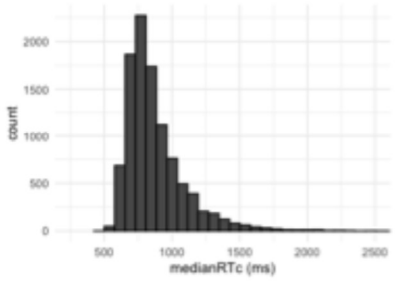
Primary Outcome:
The suggested primary outcome of the test is median reaction time of correct responses to test probes (medianRTc), a measure of processing speed. Researchers may also consider incorporating response accuracy (proportion of probes correctly matched).

User Input:
Participants respond either by touching the digit that corresponds with each symbol probe (touch compatible devices), or by pressing the corresponding digit on their keyboard.

Alternate Task Versions: Alternate forms of the test are available for repeated administration. Additionally, an ultra-brief, EMA-compatible version is available (Hawks et al., 2023; Singh et al., 2023), which has 30 seconds of test probes instead of 90 seconds.

Psychometrics:

- **Reliability**: In single-session testing, variation in performance (medianRTc) between participants has a split-half reliability of .98. See Hawks et al., 2023 and Singh et al., 2023 for psychometric details of multiple-session EMA administration.
- **Score distribution**:



References:

Chaytor, N. S., Barbosa-Leiker, C., Germine, L. T., Fonseca, L. M., McPherson, S. M., & Tuttle, K. R. (2021). Construct validity, ecological validity and acceptance of self-administered online neuropsychological assessment in adults. *The Clinical Neuropsychologist*, *35*(1), 148-164.

D'Ardenne, K., Savage, C. R., Small, D., Vainik, U., & Stoeckel, L. E. (2020). Core neuropsychological measures for obesity and diabetes trials: Initial report. *Frontiers in Psychology, 11*, 554127.

Hartshorne, J., & Germine, L. (2015) When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the lifespan. *Psychological Science,* 26(4), 433-443.

Hawks, Z. W., Strong, R., Jung, L., Beck, E. D., Passell, E. J., Grinspoon, E., ... & Germine, L. T. (2023). Accurate prediction of momentary cognition from intensive longitudinal data. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *8*(8), 841-851.

Pozo, E., T. Germine, L., Scheuer, L., & Strong, R. W. (2022). Evaluating the Reliability and Validity of the Famous Faces Doppelgangers Test, a Novel Measure of Familiar Face Recognition. *Assessment, 30*(4), 1200-1210.

Singh, S., Strong, R. W., Jung, L., Li, F. H., Grinspoon, L., Scheuer, L. S., Passell, E. J., Martini, P., Chaytor, N., Soble, J. R., & Germine, L. (2021). The TestMyBrain Digital Neuropsychology Toolkit: Development and Psychometric Characteristics. *Journal of Clinical and Experimental Neuropsychology*, *43*(8), 786-795.

Singh, S., Strong, R., Xu, I., Fonseca, L. M., Hawks, Z., Grinspoon, E., Jung, L., Li, F, Weinstock, R. S., Sliwinski, M., Chaytor, N., & Germine, L. T. (2023). Ecological Momentary Assessment of Cognition in Clinical and Community Samples: Reliability and Validity Study. *Journal of Medical Internet Research*, *25*, e45028.

Wechsler, D. (1997). WAIS-III, WMS-III Technical Manual. San Antonio, TX: The Psychological Corporation.

Appendix C

Cognitive Task Overview: TestMyBrain GradCPT

Contact: Info@ManyBrains.net
ManyBrains.net
TestMyBrain.org

TMB Test Name: TestMyBrain Gradual Onset Continuous Performance Test
Test Demo: 150 trial version
Document Version: June.05.2024
Acknowledgements:

- Original in-lab test created by: Mike Esterman, PhD and Joe DeGutis, PhD, at the Boston Attention and Learning Lab (https://www.bu.edu/ballab/index.html)
- "TMB GradCPT City/Mountain Images" by Esterman & DeGutis is licensed under CC-BY-SA 4.0.

The Many Brains Project

The Many Brains Project, is a 501(c)(3) non-profit that supports TestMyBrain (TMB) in collaboration with the Laboratory for Brain and Cognitive Health Technology at McLean Hospital and Harvard Medical School. We currently support many different types of research studies through our infrastructure for cognitive assessment - these range in size from small lab-based pilot studies to large longitudinal, multisite clinical research studies with tens of thousands of participants. As TestMyBrain has been continuously in operation since 2008, we provide a stable and secure platform for hosting and delivering mobile and web-based cognitive assessment protocols. Through TestMyBrain.org, data have been collected from over 3.7 million participants in a *citizen science* framework that includes structured return of research results toward the development, validation, and normative characterization of cognitive measures. We currently support research and education at over 2,000 sites worldwide engaged in digital neuropsychological assessment.

CITATION

Please credit The Many Brains Project and TestMyBrain in any papers, posters, or publications related to the TMB tests or data collected by TMB tests.

- Example:
  - All tasks were selected from and hosted on The Many Brains Project's web-based cognitive testing platform, TestMyBrain (Germine et al., 2012; The Many Brains Project).
    - Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, *19*(5), 847-857.
    - The Many Brains Project. *TestMyBrain Cognitive Tests*. URL: www.manybrains.net

Test Overview

Background:
The TestMyBrain Gradual Onset Continuous Performance Test (GradCPT; Fortenbaugh et al., 2015; Riley et al., 2016, 2017; Singh et al., 2021; Treviño et al., 2021; Vogel et al., 2020), adapted from Esterman et al. (2013) and Rosenberg et al. (2013), is a test of sustained attention and inhibitory control, a core aspect of executive functioning.

Task Parameters:
Participants view a sequence of 300 circular, grayscale images of cities and mountains constructed to have equal mean luminance and contrast, with the images gradually transitioning from one to the next every 800 ms. Participants are instructed to make a keyboard press (keyboard input) or tap the screen (touch input) when images of cities are presented (89.3% of trials), and withhold from responding when mountain images are presented (10.7% of trials). The total duration of the image sequence is 4 minutes. Participants complete 36 unscored practice trials before beginning the 300 test trials. All participants view the same images in the same order. In the 150-trial version of the test, participants view 133 images of cities and 17 images of mountains.

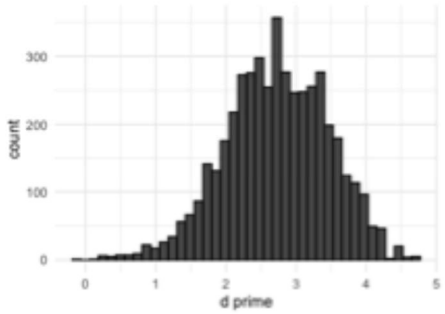
Primary Outcome:
The test's suggested primary outcome is discrimination ability (d prime), a measure of response accuracy unaffected by response bias, which represents the participant's ability to withhold responses to mountains while making responses to more prevalent city images. Secondary outcome measures of interest include response bias, commission error rate, reaction time, and reaction time variability. See Stanislaw & Todorov (1999) and the supplement of Fortenbaugh et al. (2015) for instructions on calculating d prime and other outcomes of interest.

User Input:
Participants using a keyboard are instructed to use the spacebar when making responses, but the test will accept responses from any key. If the participant's device has touch capability, at the beginning of the test the participant will be given the option to respond using touch input or keyboard input. Otherwise, the participant must use keyboard input.

Alternate Task Versions: Alternate forms of the test are available for repeated administration. Additionally, shorter versions are available, including an ultra-brief, EMA-compatible version (Hawks et al., 2023; Singh et al., 2023).

Psychometrics:

- **Reliability**: In single-session testing, variation in performance (dprime) between participants has a split-half reliability of .86. See Hawks et al., 2023 and Singh et al., 2023 for psychometric details of multiple-session EMA administration.
- **Score distribution**:



References:

Esterman, M., Noonan, S. K., Rosenberg, M., & DeGutis, J. (2013). In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. *Cerebral Cortex*, *23*(11), 2712-2723

Fortenbaugh, F. C., DeGutis, J., Germine, L., Wilmer, J. B., Grosso, M., Russo, K., & Esterman, M. (2015). Sustained Attention Across the Life Span in a Sample of 10,000: Dissociating Ability and Strategy. *Psychological Science*, *26*(9), 1497–1510.

Hawks, Z. W., Strong, R., Jung, L., Beck, E. D., Passell, E. J., Grinspoon, E., ... & Germine, L. T. (2023). Accurate prediction of momentary cognition from intensive longitudinal data. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *8*(8), 841-851.

Riley, E., Okabe, H., Germine, L., Wilmer, J., Esterman, M., & DeGutis, J. (2016). Gender differences in sustained attentional control relate to gender inequality across countries. *PloS One*, *11*(11), e0165100.

Riley, E., Esterman, M., Fortenbaugh, F. C., & DeGutis, J. (2017). Time-of-day variation in sustained attentional control. *Chronobiology International*, *34*(7), 993-1001.

Rosenberg, M., Noonan, S., DeGutis, J., & Esterman, M. (2013). Sustaining visual attention in the face of distraction: a novel gradual-onset continuous performance task. *Attention, Perception, & Psychophysics, 75*(3), 426-439.

Singh, S., Strong, R. W., Jung, L., Li, F. H., Grinspoon, L., Scheuer, L. S., Passell, E. J., Martini, P., Chaytor, N., Soble, J. R., & Germine, L. (2021). The TestMyBrain Digital Neuropsychology Toolkit: Development and Psychometric Characteristics. *Journal of Clinical and Experimental Neuropsychology*, *43*(8), 786-795.

Singh, S., Strong, R., Xu, I., Fonseca, L. M., Hawks, Z., Grinspoon, E., Jung, L., Li, F, Weinstock, R. S., Sliwinski, M., Chaytor, N., & Germine, L. T. (2023). Ecological Momentary Assessment of Cognition in Clinical and Community Samples: Reliability and Validity Study. *Journal of Medical Internet Research*, *25*, e45028

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137-149.

Treviño, M., Zhu, X., Lu, Y. Y., Scheuer, L. S., Passell, E., Huang, G. C., ... & Horowitz, T. S. (2021). How do we measure attention? Using factor analysis to establish construct validity of neuropsychological tests. *Cognitive Research: Principles and Implications*, *6*(1), 1-26.

Vogel, S. C., Esterman, M., DeGutis, J., Wilmer, J. B., Ressler, K. J., & Germine, L. T. (2020).

Childhood adversity and dimensional variations in adult sustained attention. *Frontiers in Psychology*, 11, 691.

Appendix D

Cognitive Task Overview: TestMyBrain Digit Span

Contact: Info@ManyBrains.net
ManyBrains.net
TestMyBrain.org

TMB Test Name: TestMyBrain Digit Span
Test Demo: forward version | backward version
Document Version: May.29.2024

The Many Brains Project

The Many Brains Project, is a 501(c)(3) non-profit that supports TestMyBrain (TMB) in collaboration with the Laboratory for Brain and Cognitive Health Technology at McLean Hospital and Harvard Medical School. We currently support many different types of research studies through our infrastructure for cognitive assessment - these range in size from small lab-based pilot studies to large longitudinal, multisite clinical research studies with tens of thousands of participants. As TestMyBrain has been continuously in operation since 2008, we provide a stable and secure platform for hosting and delivering mobile and web-based cognitive assessment protocols. Through TestMyBrain.org, data have been collected from over 3.7 million participants in a *citizen science* framework that includes structured return of research results toward the development, validation, and normative characterization of cognitive measures. We currently support research and education at over 2,000 sites worldwide engaged in digital neuropsychological assessment.

CITATION

Please credit The Many Brains Project and TestMyBrain in any papers, posters, or publications related to the TMB tests or data collected by TMB tests.

- Example:
  - All tasks were selected from and hosted on The Many Brains Project's web-based cognitive testing platform, TestMyBrain (Germine et al., 2012; The Many Brains Project).
    - Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, *19*(5), 847-857.
    - The Many Brains Project. *TestMyBrain Cognitive Tests*. URL: www.manybrains.net

Test Overview

Background:
TestMyBrain Digit Span (Chaytor et al., 2021; Hartshorne & Germine, 2015; Singh et al., 2021; Treviño et al., 2021) is a digital version of in-person digit span tasks (e.g., Richardson, 2007) adapted for remote administration.

Task Parameters:
After being presented visually with a set of numbers (e.g., 153), participants are asked to recall those numbers either in order (forward version, 153) or reverse order (backward version, 351). Individual numbers within each set are presented on the screen sequentially for 1000 ms each. Initially, number sets contain only two numbers, but increase up to a maximum length of 11 numbers. Participants are presented with two trials at each set length; if at least one of those two number sets is recalled successfully, the set length is increased by one number. Participants complete two unscored practice trials before beginning test trials.

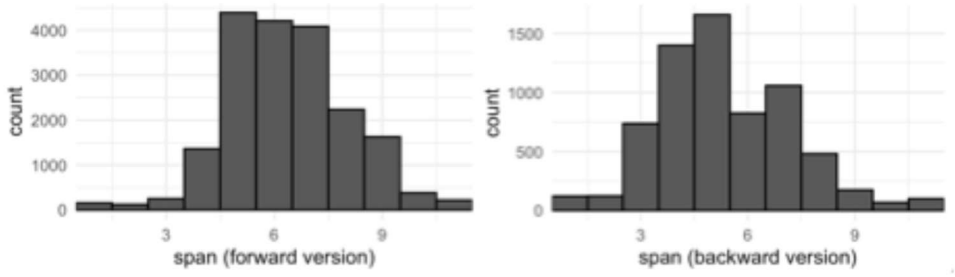
Primary Outcome:
The suggested primary outcome of the test is the length of the longest number set for which participants answered at least one trial correctly (span).

User Input:
Participants using a keyboard use the number keys to make their responses, whereas participants without a keyboard press number buttons on the screen. If the participant's device has touch capability, at the beginning of the test the participant will be given the option to respond using touch input or keyboard input. Otherwise, the participant must use keyboard input.

Alternate Task Versions: Alternate forms of the test are available for repeated administration.

Psychometrics:

- **Reliability**: see Chaytor et al. (2021)
- **Score distribution**:



References:

Chaytor, N. S., Barbosa-Leiker, C., Germine, L. T., Fonseca, L. M., McPherson, S. M., & Tuttle, K. R. (2021). Construct validity, ecological validity and acceptance of self-administered online neuropsychological assessment in adults. *The Clinical Neuropsychologist, 35*(1), 148-164.

Hartshorne, J., & Germine, L. (2015) When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the lifespan. *Psychological Science, 26*(4), 433-443.

Richardson, J. T. (2007). Measures of short-term memory: a historical review. *Cortex 43*(5), 635–650.

Singh, S., Strong, R. W., Jung, L., Li, F. H., Grinspoon, L., Scheuer, L. S., Passell, E. J., Martini, P., Chaytor, N., Soble, J. R., & Germine, L. (2021). The TestMyBrain Digital Neuropsychology Toolkit: Development and Psychometric Characteristics. *Journal of Clinical and Experimental Neuropsychology, 43*(8), 786-795.

Treviño, M., Zhu, X., Lu, Y. Y., Scheuer, L. S., Passell, E., Huang, G. C., ... & Horowitz, T. S. (2021). How do we measure attention? Using factor analysis to establish construct validity of neuropsychological tests. *Cognitive Research: Principles and Implications, 6*(1), 1-26.