Journal of the International Neuropsychological Society (2025), **31**, 242–253
doi:10.1017/S1355617725000177

Research Article

Development and validation of the TabCAT-EXAMINER: A tablet-based executive functioning battery for research and clinical trials

Mark Sanderson-Cimino[1], Katherine L. Possin[1,2], Dan M. Mungas[2,3], Emily W. Paolillo[1], Breton M. Asken[4], Elena Tsoy[1], Sabrina Jarrott[1], Yann Cobigo[1], Rowan Saloner[1], Kaitlin B. Casaletto[1], Ciaran Considine[5], Julie A. Fields[6], Joie Molden[7], Katya Rascovsky[8], Sandra Weintraub[9], Bonnie Wong[10], Hilary W. Heuer[1], Leah K. Forsberg[11], Julio C. Rojas[1], Lawren VandeVrede[1], Peter Ljubenkov[1], Gil D. Rabinovici[1], Maria Luisa Gorno-Tempini[1], William W. Seeley[1], Bruce L. Miller[1,2], Bradley F. Boeve[9], Howard J. Rosen[1,2], Adam L. Boxer[1], Katherine P. Rankin[1], Joel H. Kramer[1,2] and Adam M. Staffaroni[1]

[1]Memory and Aging Center, UCSF Weill Institute for Neurosciences, University of California San Francisco, San Francisco, CA, USA, [2]Global Brain Health Institute, University of California San Francisco, San Francisco, CA, USA, [3]Alzheimer's Disease Research Center, Department of Neurology, University of California Davis School of Medicine, Sacramento, CA, USA, [4]Florida Alzheimer's Disease Research Center, Department of Clinical and Health Psychology, University of Florida, Gainesville, Florida, USA, [5]Department of Neurology, Vanderbilt University Medical Center, Nashville, TN, USA, [6]Department of Psychiatry & Psychology, Mayo Clinic, Rochester, MN, USA, [7]Department of Neurosurgery, University of Colorado Anschutz Medical Campus School of Medicine, Aurora, CO, USA, [8]Department of Neurology and Penn Frontotemporal Degeneration Center, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA, [9]Mesulam Center for Cognitive Neurology and Alzheimer's Disease, Department of Neurology, Northwestern University, Chicago, IL, USA, [10]Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA and [11]Department of Neurology, Mayo Clinic, Rochester, MN, USA

Abstract

Objective: The National Institutes of Health (NIH) Executive Abilities: Measures and Instruments for Neurobehavioral Evaluation and Research (EXAMINER) is a validated laptop-based battery of executive functioning tests. A modified tablet version of the EXAMINER was developed on the UCSF Tablet-based Cognitive Assessment Tool (TabCAT-EXAMINER). Here we describe the battery and investigate the reliability and validity of a composite score. **Methods:** A diagnostically heterogeneous sample of 2135 individuals (mean age = 65.58, SD = 16.07), including controls and participants with a variety of neurodegenerative syndromes, completed the TabCAT-EXAMINER. A composite score was developed using confirmatory factor analysis and item response theory. Validity was evaluated via linear regressions that tested associations with neuropsychological tests, demographics, clinical diagnosis, and disease severity. Replicability of cross-sectional results was tested in a separate sample of participants ($n = 342$) recruited from a frontotemporal dementia study. As this separate sample also collected longitudinal TabCAT-EXAMINER measures, we additionally assessed test-retest reliability and associations between baseline disease severity and changes in TabCAT-EXAMINER scores. **Results:** The TabCAT-EXAMINER score was normally distributed, demonstrated high test-retest reliability, and was associated in the expected directions with independent tests of executive functioning, demographics, disease severity, and diagnosis. Greater baseline disease severity was associated with more rapid longitudinal TabCAT-EXAMINER decline. **Conclusions:** The TabCAT-EXAMINER is a tablet-based executive functioning battery developed for observational research and clinical trials. Performance can be summarized as a single composite score, and results of this study support its reliability and validity in cognitive aging and neurodegenerative disease cohorts.

Keywords: Neuropsychology; Alzheimer's disease; frontotemporal dementia; digital; computerized cognitive assessment

(Received 22 October 2024; final revision 20 February 2025; accepted 3 April 2025)

Research Significance

Research Question(s) or Topic(s): The accurate and efficient measurement of executive functioning is important for a range of clinical conditions including neurodegenerative diseases such as Alzheimer's disease and frontotemporal dementia. Tablet-based assessments provide many advantages over traditional and computer-based testing. A highly utilized tablet-based cognitive battery, TabCAT, includes measures of executive functioning,

Corresponding authors: Mark Sanderson-Cimino; Email: mark.sandersoncimino@ucsf.edu, Adam M. Staffaroni, email: adam.staffaroni@ucsf.edu
Joel Kramer and Adam Staffaroni are co-senior authors.

Cite this article: Sanderson-Cimino M., Possin K.L., Mungas D.M., Paolillo E.W., Asken B.M., Tsoy E., Jarrott S., Cobigo Y., Saloner R., Casaletto K.B., Considine C., Fields J.A., Molden J., Rascovsky K., Weintraub S., Wong B., Heuer H.W., Forsberg L.K., Rojas J.C., VandeVrede L., Ljubenkov P., Rabinovici G.D., Gorno-Tempini M.L., Seeley W.W., Miller B.L., Boeve B.F., Rosen H.J., Boxer A.L., Rankin K.P., Kramer J.H., & Staffaroni A.M. (2025) Development and validation of the TabCAT-EXAMINER: A tablet-based executive functioning battery for research and clinical trials. *Journal of the International Neuropsychological Society*, **31**: 242–253, https://doi.org/10.1017/S1355617725000177

© The Author(s), 2025. Published by Cambridge University Press on behalf of International Neuropsychological Society

https://doi.org/10.1017/S1355617725000177 Published online by Cambridge University Press

but lacks a single outcome measure. Here we develop and validate a single tablet-based summary measure of executive functioning. **Main Findings:** Modern statistical procedures were used to create an executive functioning composite score. We show that this score is reliable and associated as expected with other cognitive tasks. We also demonstrate that participants with greater clinical severity perform worse on this measure than individuals with less impairment and show faster rates of longitudinal decline. **Study Contributions:** The TabCAT-EXAMINER will aid assessment of executive functioning in observational research and clinical trials.

Introduction

Executive functioning is an umbrella term for a multidimensional set of cognitive processes that influence goal-directed behavior (Karr et al., 2018; Rabinovici et al., 2015). It is supported by a dispersed but connected set of brain regions (Alvarez & Emory, 2006; Bettcher et al., 2016; Kramer et al., 2014; Niendam et al., 2012; Reineberg & Banich, 2016) and is impacted by neurological conditions including epilepsy (Arrotta et al., 2022; Longo et al., 2013), multiple sclerosis (Foong et al., 1997; Roman & Arnett, 2016), traumatic brain injury (Azouvi et al., 2017; Broadway et al., 2019), and neurodegenerative diseases such as Alzheimer's disease and frontotemporal dementia (FTD) (Guarino et al., 2019; Rabinovici et al., 2015; Staffaroni, Bajorek, et al., 2020; Stopford et al., 2012). The parcellation of executive functioning into subdomains varies by study, but cognitive neuroscience and statistical approaches commonly identify subdomains such as inhibition, set-shifting, fluency/generativity, and working memory (Kramer, 2014; Miyake et al., 2000). The assessment of executive functioning is complicated by the breadth of cognitive tests used to measure these domains, many of which focus on a specific subcomponent (e.g., inhibition) (Faria et al., 2015). As a result, clinics and research studies differ in what aspect of executive functioning they are describing, which may attenuate reproducibility and hamper cross-study comparisons. To address these challenges, the National Institute of Health (NIH) funded the Executive Abilities: Measures and Instruments for Neurobehavioral Evaluation and Research (EXAMINER) project to develop a laptop-based battery of executive functioning tests that was informed by cognitive neuroscience, and could enable measurement of executive functioning in clinical trials and observational research (Kramer et al., 2014; Kramer, 2014). The project aimed to create a modular, flexible battery that captured the subcomponents of executive functioning. NIH-EXAMINER tasks of working memory, verbal fluency, and cognitive control were modeled using confirmatory factor analysis and item response theory (IRT) to create an Executive Composite score that summarized executive functioning into a single pragmatic and psychometrically robust outcome (Bott et al., 2014; Possin et al., 2014).

Over the last decade, the NIH-EXAMINER and the Executive Composite score have been implemented and validated in myriad studies. For example, scores on the NIH-EXAMINER are strongly associated with real-world executive behavior and with volumes of frontal, parietal, and subcortical brain regions known to support executive functioning (Possin et al., 2014). The NIH-EXAMINER has also shown utility for detecting and monitoring symptoms of neurological conditions including traumatic brain injury (Broadway et al., 2019), attention deficit disorder (Schreiber et al., 2014), and pediatric sickle cell disease (Schatz et al., 2014). The NIH-EXAMINER is sensitive to cognitive changes across a range of neurodegenerative diseases (You et al., 2014), and worse performance has been associated with greater disease severity, suggesting it measures clinically meaningful symptoms (Possin et al., 2014; Staffaroni, Bajorek, et al., 2020). Moreover, the Executive Composite score is a sensitive marker for early disease manifestations, as demonstrated by its ability to detect cross-sectional and longitudinal changes in presymptomatic FTD mutation carriers (Staffaroni, Bajorek, et al., 2020). The NIH-EXAMINER has been used to measure treatment effects on cognitive performance in pharmacological treatment trials (Assogna et al., 2021; Leite et al., 2022; Miller et al., 2014; Yu et al., 2018), including those targeting AD and other causes of mild cognitive impairment and mild dementia (Hajjar et al., 2020; Vossel et al., 2021), as well as non-pharmacological trials of mindfulness, cognitive training, and transcranial direct current stimulation (Assogna et al., 2021; Clark et al., 2021; Leite et al., 2022; Yu et al., 2018).

Since the development of the laptop-based NIH-EXAMINER, tablets have become a frequently used alternative to paper-and-pencil testing or laptops for cognitive assessments. Tablets offer similar benefits as computer-based testing (e.g., accurate logging of reaction time, automated scoring), but offer additional advantages with respect to portability, affordability, and user friendliness (Canini et al., 2014; Vaportzis et al., 2017; Wilson et al., 2022). Staffaroni et al., (2020) and Wilson et al., (2022) provide brief reviews on the benefits of tablet-based assessments over computer-based assessments. One such tablet-based testing platform is the UCSF Tablet-based Cognitive Assessment Tool (TabCAT; https://tabcathealth.com/). A subset of the validated NIH-EXAMINER tasks was programmed into the TabCAT platform (TabCAT-EXAMINER). Based on user feedback and data analyses of original NIH-EXAMINER tasks, slight modifications were made to some tasks (e.g., reducing the number of practice trials), and a new task, Running Dots, was added in place of the N-Back task (see Methods: TabCAT-EXAMINER). Prior studies have demonstrated that changing the administration platform from a computer to tablet can impact scores. For example, the NIH Toolbox is a set of cognitive tasks that was adapted from computer to tablet, with subtle alterations to the tasks. These alterations led to critical differences in performance for the majority of tasks, necessitating the re-creation of validation and normative data (Brearly et al., 2018; Germine et al., 2019). We therefore set out to recreate and validate an EXAMINER executive functioning composite for tablets.

Here we describe the development and validation of the TabCAT-EXAMINER score, a measure of executive functioning. The precision of a composite score across a range of clinical presentations is improved when that score is developed in a heterogenous testing cohort (Kramer et al., 2014). As such, model development and initial validation analyses were conducted in a large, diagnostically heterogenous cohort of individuals with AD and related disorders, including FTD. Given the high frequency of executive dysfunction in FTD and the reported usefulness of the original NIH-EXAMINER in monitoring FTD (Miller et al., 2014), we selected participants from the ARTFL-LEFFTDS Longitudinal Frontotemporal Lobar Degeneration (ALLFTD) consortium with longitudinal TabCAT-EXAMINER data as a validation cohort. In addition to replicating findings observed in the development cohort, we conducted analyses in this validation cohort to further understand longitudinal performances, including test-retest reliability. We hypothesized that a single, general TabCAT-EXAMINER score will 1) fit the data well; 2) demonstrate

https://doi.org/10.1017/S1355617725000177 Published online by Cambridge University Press

convergent and divergent validity with other independent cognitive tasks; 3) be sensitive to a variety of neurodegenerative syndromes and associated with worsening disease severity; and 4) be sensitive to subtle cognitive changes in presymptomatic carriers of autosomal dominant FTD mutations.

Materials and Methods

Participants

All participants completed at least one of the five tablet-based TabCAT-EXAMINER measures. Some participants were previously exposed to the verbal fluency measures which are administered as part of the standard Uniform Data Set (UDS) Version 3.0 neuropsychological battery, but baseline data represent the first time that participants completed one of the TabCAT-EXAMINER measures. Inclusion criteria required English as the primary language to remove variance associated with language that may impact performance on these tasks. An associated study is underway to evaluate the TabCAT-EXAMINER in a Spanish speaking cohort (see future directions). Written informed consent was obtained for all participants or their legal representative. Studies were approved by the Institutional Review Boards at University of California, San Francisco (UCSF) and Johns Hopkins University (Trial Innovation Network) in accordance with university institutional guidelines and the Helsinki Declaration.

The total sample comprised 2135 participants. A subset of these participants were used as the developmental sample ($n = 1793$; Table 1), which included individuals recruited through UCSF ($n = 509$) and the ALLFTD research consortium ($n = 1284$), a multi-center study of FTD (Boeve et al., 2020). Similar to the original NIH-EXAMINER development sample (Kramer et al., 2014; Possin et al., 2014), this development subsample was symptomatically heterogeneous and syndromic diagnoses were made in consensus conferences using published criteria (Barker et al., 2022; Brooks et al., 2000; Dubois et al., 2007; Gorno-Tempini et al., 2011; Jack Jr et al., 2018; Litvan et al., 1998; McKeith et al., 2017; Neary et al., 1998; Petersen et al., 2010; Petersen et al., 2013; Petersen et al., 2014; Postuma et al., 2015; Rascovsky et al., 2011; Sachdev et al., 2014). The sample included participants from UCSF and ALLFTD diagnosed with/as cognitively unimpaired (CU; $n = 786$), mild cognitive impairment (MCI; $n = 237$), Alzheimer's dementia (AD; $n = 115$), behavioral variant FTD (bvFTD; $n = 167$), corticobasal syndrome (CBS; $n = 29$), logopenic variant primary progressive aphasia (lvPPA; $n = 28$), nonfluent variant PPA (nvPPA; $n = 44$), progressive supranuclear palsy (PSP; $n = 37$), and semantic variant PPA (svPPA; $n = 49$). There were 112 participants with non-neurodegenerative diagnoses (e.g., primary mood etiologies) and/or small subsamples (e.g., amyotrophic lateral sclerosis: $n < 10$) labeled as "other." There were 184 participants who were missing diagnosis. The MCI group included those with mild cognitive deficits and/or mild behavioral changes irrespective of suspected etiology (Barker et al., 2022).

A different subset of ALLFTD participants ($n = 342$) not included in the development cohort comprised the validation cohort. This subsample had longitudinal TabCAT-EXAMINER data, completed 6–24 months after their baseline (mean follow-up length: 368 days; SD = 64). This sample included those diagnosed as: CU ($n = 175$), MCI ($n = 26$), AD ($n = 10$), bvFTD ($n = 46$), CBS ($n = 5$), lvPPA ($n = 1$), nvPPA ($n = 15$), PSP ($n = 8$), and svPPA ($n = 14$). There were 127 participants who fell into the categories labeled as "other" in the development sample (e.g., primary etiology mood, $n = 3$; Parkinson's disease, $n = 1$). There were 85 participants who were missing diagnosis.

A description of the sample with genetic testing is presented in Supplemental Tables 1 and 2. DNA samples were screened using targeted sequencing of a custom panel of genes previously implicated in neurodegenerative diseases, including *GRN* and *MAPT*. Hexanucleotide repeat expansions in *C9orf72* were detected using both fluorescent and repeat-primed PCR (Ramos et al., 2020). The developmental cohort included 79 participants who carried pathogenic mutations (*MAPT*, $n = 23$; *C9orf72*, $n = 43$; and *GRN*, $n = 13$) known to cause familial frontotemporal lobar degeneration (f-FTLD). There were also 132 non-carriers from families with an identified pathogenic f-FTLD mutation (i.e., family controls). The validation cohort included 80 participants who carried f-FTLD genetic mutations (*MAPT*, $n = 24$; *C9orf72*, $n = 44$; and *GRN*, $n = 12$). There were also 101 non-carriers from families with an identified pathogenic f-FTLD mutation in the validation cohort.

TabCAT-EXAMINER

The TabCAT-EXAMINER comprises nine executive functioning tasks, including five tablet-based measures and four measures of verbal fluency. Tablet-based measures included tasks of inhibition (Flanker), set-shifting (Set Shifting), processing speed/working memory (Match), and working memory (Dot Counting, Running Dots) (Mullady et al., 2022; Possin et al., 2018; Weintraub et al., 2009). These tasks are also influenced by processing speed. Running Dots replaced the N-back task in the NIH-EXAMINER in response to user feedback. Match, a brief, well-validated assessment of executive functioning was also added to the battery to expand upon tasks modified from the NIH-EXAMINER. Other differences between the TabCAT-EXAMINER and the original include alterations due to the change in format (computer vs tablet) such as screen size, instructions, stimuli (i.e., size and shape) and user interface. Additional task-specific alterations are described below and in more detail at https://tabcathealth.com/. Images of example TabCAT tasks are presented in Supplemental Figure 1.

Flanker

The participant is presented with a display of arrows and asked to indicate which direction the central arrow is pointing. There are two conditions: in the congruent condition the non-target stimuli point in the same direction as the central arrow; in the non-congruent condition, the non-target stimuli point in the opposite direction of the central arrow. This task was modeled after the original NIH-EXAMINER Flanker task (Kramer et al., 2014). Alterations included a reduction in the total number of practice and response trials, as well as slight changes in the minimum and maximum response times allowed. The primary outcome for this task is a score that combines speed and accuracy (range: 0–10); see (Mullady et al., 2022; Kramer et al., 2014; Weintraub et al., 2013) for details.

Set shifting

Participants are asked to match a stimulus presented at the top of the screen to one of two stimuli at the bottom of the screen. In the task-homogenous condition, participants are asked to match on shape or color. In the task-heterogenous condition, participants are asked to alternate between matching color and shape, pseudo-randomly. Similar to Flanker, this task was modeled after the

https://doi.org/10.1017/S1355617725000177 Published online by Cambridge University Press

Table 1. Sample description and TabCAT-EXAMINER scores by CDR®+NACC-FTLD Global Score subgroups

		CDR®+NACC-FTLD Global Score*					
	Full sample	**0**	**0.5**	**1**	**2**	**3**	**Comparison**
Demographics							
Sample size	1793	656	326	302	130	10	–
Age	64.22 (13.94)	59.19 (16.60)	66.87 (11.23)	66.65 (8.92)	66.49 (9.58)	71.00 (8.76)	0 < 0.5
Education	16.38 (2.77)	16.84 (2.36)	16.42 (2.78)	15.86 (2.81)	15.59 (2.91)	16.38 (2.77)	0 > 0.5 > 1, 2
Male *n* (%)	860 (48.4)	274 (41.8)	172 (52.8)	183 (60.6)	78 (60.0)	3 (30.0)	0 < 0.5 < 1, 2
Diagnosis *n* (%)							
CU	786 (49.0)	573 (91.2)	56 (18.1)	12 (4.0)	6 (4.7)	–	
MCI	237 (14.8)	20 (3.2)	147 (47.4)	26 (8.8)	4 (3.1)	–	
AD	115 (7.2)	–	11 (3.5)	72 (24.2)	12 (9.4)	–	
CBS	29 (1.8)	–	9 (2.9)	12 (4.0)	4 (3.1)	1 (10.0)	
lvPPA	28 (1.7)	–	3 (1.0)	19 (6.4)	6 (4.7)	–	
bvFTD	167 (10.4)	–	25 (8.1)	65 (21.9)	62 (48.4)	9 (90.0)	
nvPPA	44 (2.7)	–	12 (3.9)	23 (7.7)	6 (4.7)	–	
PSP	37 (2.3)	–	5 (1.6)	18 (6.1)	10 (7.8)	–	
svPPA	49 (3.1)	–	13 (4.2)	29 (9.8)	6 (4.7)	–	
Other	112 (7.0)	35 (5.6)	29 (9.4)	21 (7.1)	12 (9.4)	–	
TabCAT-EXAMINER							
Factor score	−0.03 (0.91)	0.56 (0.61)	−0.19 (0.70)	−0.75 (0.73)	−1.06 (0.76)	−1.91 (0.62)	0 > 0.5 > 1 > 2
Flanker	7.37 (1.54)	8.09 (1.00)	7.05 (1.41)	6.35 (1.72)	5.66 (1.55)	5.46 (0.30)	0 > 0.5 > 1
Set shifting	7.04 (1.52)	7.84 (1.02)	6.56 (1.37)	5.93 (1.70)	5.29 (1.31)	5.25 (1.21)	0 > 0.5 > 1
Match	41.91 (13.80)	50.71 (7.64)	40.62 (10.41)	31.12 (13.13)	26.27 (14.04)	22.70 (15.04)	0 > 0.5 > 1 > 2
Dot counting	15.23 (5.80)	17.92 (4.54)	14.51 (4.95)	11.88 (5.41)	9.80 (7.07)	3.50 (3.15)	0 > 0.5 > 1
Running dots	29.65 (7.86)	32.40 (6.77)	26.38 (7.46)	28.15 (8.11)	25.82 (7.68)	29.00 (–)	0.5 > 1
Animals	18.63 (8.97)	24.90 (6.17)	17.05 (7.01)	12.31 (6.57)	10.00 (5.87)	3.75 (3.11)	0 > 0.5 > 1
Vegetables	11.90 (6.25)	16.04 (4.69)	10.41 (4.57)	7.90 (5.19)	6.13 (4.31)	2.88 (2.10)	0 > 0.5 > 1
Words (F)	12.36 (6.43)	16.26 (4.60)	11.46 (5.41)	8.25 (5.54)	6.71 (4.88)	2.38 (1.77)	0 > 0.5 > 1
Words (L)	11.44 (5.78)	14.69 (4.13)	10.93 (4.90)	8.02 (5.14)	6.20 (5.08)	2.75 (1.49)	0 > 0.5 > 1

Note: The "Comparison" column indicates significance of comparisons (T-test or Chi^2 test) between CDR®+NACC-FTLD Global score groups. Statistical comparisons involving the CDR®+NACC-FTLD Global Score = 3 group were not computed due to small sample size. Means, (standard deviations), and, when applicable, sample size [n], are presented. There were 368 participants missing CDR®+NACC-FTLD Global Scores.
CU = cognitively unimpaired, MCI = mild cognitive impairment, AD = Alzheimer's dementia, bvFTD = behavioral variant Frontotemporal dementia, CBS = corticobasal syndrome, lvPPA = logopenic primary progressive aphasia, nvPPA = nonfluent variant PPA, PSP = progressive supranuclear palsy, svPPA = semantic variant PPA.
Other: participants with non-neurodegenerative diagnoses (e.g., primary mood etiologies) and/or small subsamples (e.g., amyotrophic lateral sclerosis: $n < 10$).

original NIH-EXAMINER Set Shifting task and underwent minor alterations for the TabCAT system: reduction in practice trials, slight alterations to stimuli size and shape, and a slight change in the minimum and maximum allowed response times. Similar to Flanker, the outcome variable combines accuracy and reaction time across all task-heterogenous conditions (range: 0–10) (Kramer et al., 2014).

Match

Participants are shown a set of numbers that are each matched to a simple shape in a key at the top of the screen. At the bottom of the screen the same set of shapes are presented. During each trial, a number appears at the center of the screen. Participants are asked to tap the corresponding shape at the bottom of a screen as quickly as possible for two minutes. Like the classic Digit Symbol Substitution Test paradigm on which it was based, Match is a highly sensitive but not specific measure of executive functioning (Possin et al., 2018) that relies on processing speed, attention, visual-motor coordination, and working memory (Jaeger, 2018). Although Digit Symbol Substitution Tests are affected by processing speed, performance on these tasks is also affected by working memory and other executive functioning abilities (Baudouin et al., 2009; Jaeger, 2018; Piccinin & Rabbitt, 1999; Salthouse, 1992).

Dot counting

Participants are asked to count the total number of blue circles on a display. The number of circles ranges from two to nine per display, with additional distractors (e.g., squares). The number of displays presented increases each trial, and at the end of each trial, the participant is asked to recall the total number of blue circles per display. There are six total trials. This task is modeled after the original NIH-EXAMINER Dot Counting task and assesses verbal working memory (Kramer et al., 2014). Compared to the NIH-EXAMINER, the TabCAT-EXAMINER version altered the number and size of stimuli at each trial. The outcome variable is the total number of correct responses across all trials (range: 0–27) (Mullady et al., 2022).

Running dots

Participants are shown a 16-square grid where a series of dots appear, one by one. They are asked to recall the last two to five locations where a dot appeared, after the dots disappear from the screen. The condition defining the target dots varies (e.g., the last two dots vs last four dots), as does the number of dots presented before the target dots. Participants are awarded a point for each dot they recall in the correct location and order; they are awarded half a point for recalling the correct location of a dot, but in the incorrect order. This task assessing spatial working memory was designed and implemented only in the TabCAT system. The primary outcome variable is the sum of total points across all trials (range: 0–42).

Verbal fluency

Measures included semantic fluency (Animals, Vegetables) and phonemic fluency (words [L], words [F]) collected as part of the

https://doi.org/10.1017/S1355617725000177 Published online by Cambridge University Press

UDS version 3.0 Neuropsychological Battery (Weintraub et al., 2018). Although these tasks are not administered on the tablet, the TabCAT system allows the user to input verbal fluency scores.

Other assessments

Additional cognitive tasks. In addition to the TabCAT-EXAMINER, participants completed a battery of standard paper-and-pencil neuropsychological measures within 180 days of tablet testing (supplemental methods). Measures included four UDS executive functioning tasks (Lezak et al., 2004; Possin et al., 2011; Weintraub et al., 2009). Participants also completed tasks of visuospatial ability (Benson Figure Copy), visual memory (Benson Figure Recall; Possin et al., 2011), verbal memory (Craft Story Recall (Craft et al., 1996) and California Verbal Learning Test-II, Short Form (Delis et al., 2000), and naming (Boston Naming test; Weintraub et al., 2018 or Multilingual Naming Test; Stasenko et al., 2019). Participants scores on the Multilingual Naming Test were converted to Boston Naming Test scores based on a prior crosswalk analysis (Monsell et al., 2016). Finally, participants completed two additional measures on the TabCAT platform that are not part of the EXAMINER: Line Orientation, a measure of visuospatial perception, and Favorites, a measure of associative memory (Possin et al., 2018).

Clinical Dementia Rating Scale Dementia Staging Instrument plus NACC Behavior and Language Domains (CDR®+NACC-FTLD). The CDR®+NACC-FTLD (Knopman et al., 2008) is a clinician-administered, semi-structured interview with participants and informants that adds language and behavior domains to the original CDR (Morris, 1997) to account for core impairments of FTD-spectrum disorders. An algorithm (Miyagawa et al., 2020) was used to calculate a Global score that can be used to characterize disease severity as asymptomatic (0), prodromal (0.5), or symptomatic (1, 2, or 3). Within the symptomatic group a score of 1, 2, and 3, refers to mild, moderate, and severe dementia, respectively. A sum of the eight domain Box Scores (CDR®+NACC-FTLD sum of boxes) was also calculated (max score = 24). Greater Global and Box Scores indicate more functional impairment (Miyagawa et al., 2019, 2020). This measure was administered to participants with and without FTD.

Factor score modeling

Model building was conducted using the statistical software R (Team, 2021) and is fully described in the supplement. Briefly, exploratory factor analysis was used to determine the number of factors. A confirmatory model was then fit, followed by a multilevel IRT model using the MIRT package (Chalmers, 2012).

IRT is a form of latent variable modeling, similar to confirmatory factor analysis. This method differs from traditional factor analysis in that it models item difficulty as well as the item's ability to discriminate between participants with different abilities. As a result, items in IRT models are weighted in how effectively they assess the target ability at varying levels along that ability continuum (Mungas et al., 2003). In the present analyses, we utilized a graded response model that allows for ordinal variables (Samejima, 2011) and recoded continuous variables into ordinal scores with up to 12 response categories and at least five observations in each category. This method retained the overall distribution of the raw, continuous data. The IRT model set the loadings within each fluency factor to be equal; this approach in MIRT is equivalent to specifying correlated measurement error in a CFA model. Acceptable model fit was determined via factor loadings (>.4 retained; Matsunaga, 2010; Williams et al., 2010), the Tucker Lewis Index (TLI; >0.95; Schumacker & Lomax, 2004), the Comparative Fit Index (CFI; <0.90; Savalei & Bentler, 2006), and the Root Mean Square Error of Approximation (RMSEA; <.08; Byrne, 2001).

A series of analyses were conducted to determine the reliability and utility of the TabCAT-EXAMINER score when calculated with fewer components, mimicking scenarios in which not all tests were administered. We retained the thresholds and loadings from the full model (described above) in a series of sub-models that retained two to seven components; all possible 502 combinations were evaluated. A separate factor score was generated by each model. Models were then grouped by number of components (e.g., two-component model, three-component model) and by requiring a specific component be included (e.g., two-factor models that included Match). To evaluate the reliability and internal consistency of the TabCAT-EXAMINER score, each reduced factor score was correlated with the full, nine-component TabCAT-EXAMINER. To evaluate the utility of a factor score generated with fewer than nine components, each reduced factor score was also correlated with the CDR®+NACC-FTLD Sum of Boxes. Results were separated by the number of components included in the model and by specific components.

Convergent and divergent validity

To evaluate construct validity, we tested the association of the TabCAT-EXAMINER score with independent neuropsychological measures and scales of disease severity. First, a linear regression examined the association between the CDR®+NACC-FTLD Sum of Boxes and the TabCAT-EXAMINER score. Next, separate linear regressions were fit to assess the association between the TabCAT-EXAMINER score and age, education, sex, and neuropsychological measures while controlling for the CDR®+NACC-FTLD Sum of Boxes. Correlations did not adjust for demographic variables, as the goal was to understand the relationship between the TabCAT-EXAMINER score and paper-and-pencil measures rather than to evaluate the scores incremental utility above and beyond demographics. In a second set of analyses, linear regression was used to evaluate differences in TabCAT-EXAMINER scores across levels of clinical severity (CDR®+NACC-FTLD Global score), controlling for age, education, and sex. As we would expect executive functioning to differ between syndromic diagnostic groups and controls, we then demonstrated known group validity by assessing differences in TabCAT-EXAMINER scores across diagnostic syndromes through linear regressions, controlling for age, sex, and education. Based on prior studies (Possin et al., 2014; Staffaroni et al., 2021), we hypothesized a similar degree of executive impairment across diagnostic groups. We next evaluated the ability of the TabCAT-EXAMINER to differentiate asymptomatic and prodromal f-FTLD mutation carriers (CDR®+NACC-FTLD Global score <1; $n = 33$) from family controls (CDR®+NACC-FTLD Global score = 0; $n = 56$) using regression. Covariates included education, age, and sex. This analysis was restricted to those over the age of 45 to remove participants likely to be many years from symptom onset based on natural history studies (Staffaroni et al., 2022).

Validation cohort and longitudinal analyses

Cross-sectional analyses. Mirroring analyses in the development sample, regressions assessed the association between the TabCAT-EXAMINER score and age, education, sex, CDR®+NACC-FTLD

https://doi.org/10.1017/S1355617725000177 Published online by Cambridge University Press

Sum of Boxes, and neuropsychological measures. Regressions next compared TabCAT-EXAMINER scores across diagnostic syndromes. We then replicated comparisons between asymptomatic and prodromal f-FTLD mutation carriers ($n = 37$) from family controls ($n = 31$) using regressions in a sample over the age of 45.

Longitudinal analyses. First, in a sample defined as functionally intact (CDR®+NACC-FTLD Global score = 0) at baseline and follow-up, test-retest reliability was assessed as the correlation between baseline and longitudinal scores. Second, to test the hypothesis that rates of decline in TabCAT-EXAMINER scores hasten as disease severity increases, we fitted linear regressions with TabCAT-EXAMINER annualized change scores ([visit 2 – visit 1] / retest-interval) as the outcome and baseline disease severity (CDR®+NACC-FTLD Global score group) as the predictor. Third, we fitted linear regressions to test whether those who converted to a higher CDR®+NACC-FTLD Global score at follow up showed more rapid TabCAT-EXAMINER declines than those who remained stable at follow up. Covariates included education, sex, and retest interval.

Results

Model building

The exploratory factor analysis scree plot (Supplemental Figure 2) suggested a one factor model accounted for the majority of variance (84%): $\text{eigenvalue}_{1\text{-factor}}$: 5.53; $\text{eigenvalue}_{2\text{-factor}}$: 1.0; $\text{eigenvalue}_{>2\text{-factors}}$: mean = .35 (SD = .23). The CFA model, which accounted for shared measurement variance of the verbal fluency tasks, fit the data well (CFI = 0.996, TLI = 0.995, RMSEA = 0.028) with moderate-to-strong factor loadings. This final model (Figure 1) was specified as a multilevel IRT model to generate factor scores.

Factor score distribution and reliability are detailed in the supplement (Supplemental Figures 3–5). Briefly, the TabCAT-EXAMINER score was normally distributed. A series of models calculated the TabCAT-EXAMINER with fewer than nine components and correlated these models with the full, nine component TabCAT-EXAMINER. As the number of components included in a model decreased, there was a reduction in the average correlation and an increase in uncertainty of the estimated association. However, the average correlation with the full TabCAT-EXAMINER remained strong, ranging from .84 to .90 across models. These results are presented in Supplemental Figure 4. The association of the reduced models and the full TabCAT-EXAMINER varied somewhat by specific component (Supplemental Figure 5), particularly when fewer than four components were included. There was also variability in the association between the reduced factor score and the CDR®+NACC-FTLD Sum of Boxes when fewer components were included (Supplemental Figure 6).

Developmental sample: convergent and divergent validity

Associations between the TabCAT-EXAMINER score and paper-and-pencil neuropsychological tasks in the developmental sample are presented in Figure 2 (circles). Higher TabCAT-EXAMINER scores (better performances) were associated with more years of education ($\beta = .09$, [95%CI: .05, .13]; $p < .001$; $\text{partial-eta}^2 = .07$) and younger age ($\beta = -.35$, [95%CI: −.38, −.31], $p < .001$; $\text{partial-eta}^2 = .27$). Additionally, paper-and-pencil executive functioning tasks were significantly associated with the TabCAT-EXAMINER scores in the expected direction, and standardized betas suggest large effects (β's range; .37 to .52; p's < .001; partial-eta^2 range: .45 to .59). The tasks of non-executive domains (i.e., language, memory, visuospatial) were also significantly associated with the TabCAT-EXAMINER score, with modest reductions in magnitude (β's range: .13 to .42; partial-eta^2 range: .25 to .47) compared to the magnitude of associations with tasks of executive functioning.

Lower TabCAT-EXAMINER scores were strongly correlated with greater disease severity as measured by the CDR®+NACC-FTLD Box Score ($\beta = -.65$, [95%CI: −.67, −.61], $p < .001$; $\text{partial-eta}^2 = .43$), and remained associated after adjusting for covariates ($\beta = -.16$, [95%CI: −.16, −.14], $p < .001$ $\text{partial-eta}^2 = .47$). There was a stepwise decline in factor score with increasing CDR®+NACC-FTLD Global score ($0 > 0.5 > 1 > 2$; $p < 0.001$). Table 1 provides summary scores by CDR®+NACC-FTLD Global scores across all TabCAT-EXAMINER components and factor score.

The functionally intact control group had a significantly greater TabCAT-EXAMINER score than any of the symptomatic groups (Figure 3; p's < .006). As a follow-up analysis, the control group was compared to symptomatic groups with at least 1 positive AD-biomarker. AD-biomarker positivity was based on PET-imaging, CSF, and/or plasma and diagnoses were determined in a consensus conferences using published criteria (Jack Jr et al., 2018). Results supported a stepwise decline in factor scores (p's < .001) with increasing severity: cognitively unimpaired ($n = 647$; mean = .53, SD = .63), mild cognitive impairment likely due to Alzheimer's disease (1+ AD biomarker; $n = 40$; mean = −.4; SD = .54), and Alzheimer's disease dementia (1+ AD biomarker; $n = 37$; mean = −1.13, SD = .74) groups.

In a subsample of individuals over the age of 45 ($n = 89$), the average score for asymptomatic/prodromal f-FTLD mutation carriers was worse than family controls, although this difference was not statistically significant ($\beta = -.21$, [95%CI: −.43, −.01], $p = .06$).

Validation sample and longitudinal analyses

Figure 2 provides associations among the TabCAT-EXAMINER score, demographics, and paper-and-pencil neuropsychological tasks in the ALLFTD validation sample (triangles). This figure demonstrates the similarities between standardized regression coefficients in the validation and development sample (e.g., association with Trail Making Test B: ALLFTD Validation: $\beta = .54$; Development: $\beta = .52$). The TabCAT-EXAMINER score was correlated with baseline CDR®+NACC-FTLD Box Score ($\beta = -.74$, [95%CI: −.78, −.70], $p < .001$; $\text{partial-eta}^2 = .43$), and this association remained significant while controlling for age, education, and sex, ($\beta = -.16$, [95%CI: −.18, −.14], $p < .001$ $\text{partial-eta}^2 = .65$). Similar to findings in the developmental sample, a stepwise decline in factor score was observed with increasing CDR®+NACC-FTLD Global score ($0 > 0.5 > 1 > 2$; p's < .01). As found in developmental cohort, the majority of diagnostic groups performed significantly worse on the TabCAT-EXAMINER score than the cognitively unimpaired sample (i.e., CU > CBS, bvFTD, PPA, PSP, svPPA). Details are presented in Supplemental Figure 7. Within a small sample of asymptomatic/prodromal individuals from families known to carry f-FTLD mutations ($n = 68$), those with mutations showed a trend towards worse performance on average than noncarrier family members on the baseline

https://doi.org/10.1017/S1355617725000177 Published online by Cambridge University Press

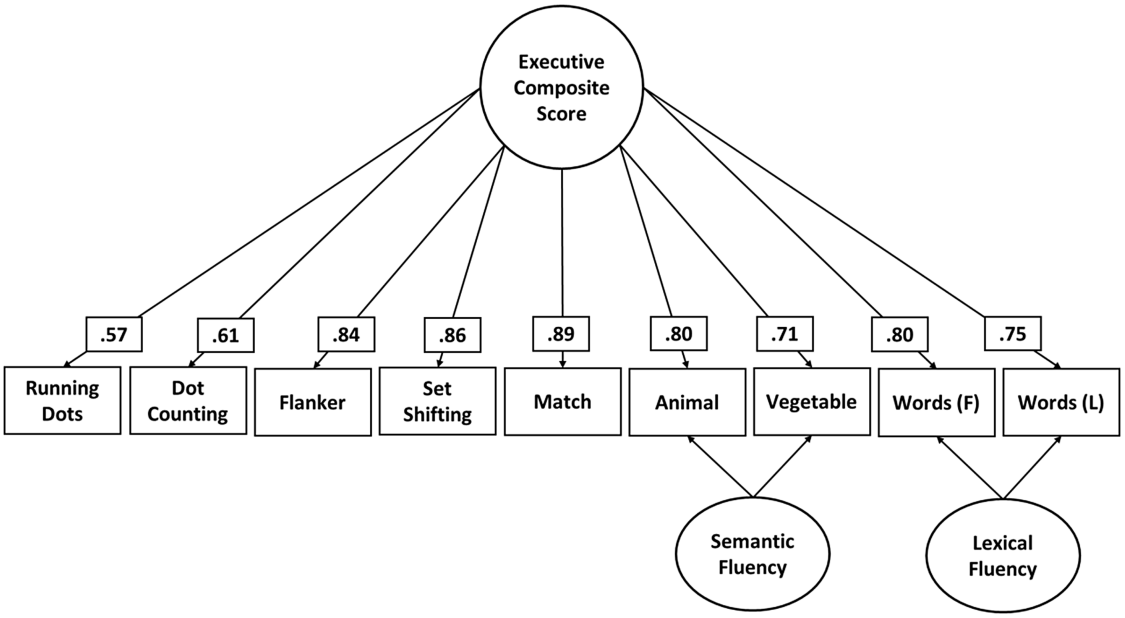


Figure 1. TabCAT-EXAMINER confirmatory factor analysis. *Note*. Standardized factor loadings are presented for the final confirmatory factor analysis model. Latent factors are represented by circles/ovals. Rectangles represent measured items.

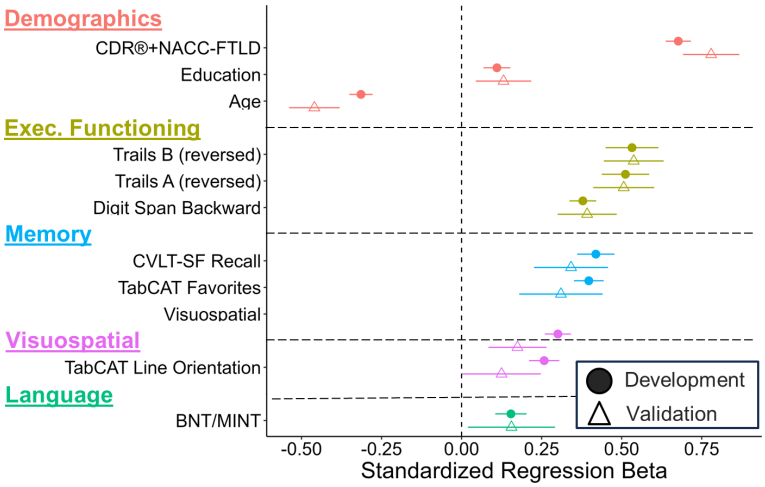


Figure 2. Associations of the TabCAT-EXAMINER score with demographics and clinical measures. Forest plot presenting standardized regression betas in the developmental (circle) and validation cohorts (triangle). The first row presents the association between CDR®+NACC-FTLD Sum of Boxes (reverse scored) and the TabCAT-EXAMINER. All other regression models control for CDR®+NACC-FTLD Sum of Boxes. Confidence intervals are shown via the horizontal lines. BNT = Boston Naming Test, MINT = Multilingual Naming Task.

TabCAT-EXAMINER ($\beta = -.17$; 95%CI: $-.38$, $.05$; $p = .13$), although this difference was not statistically significant. Of note, this association was similar in magnitude to that observed in the development sample ($\beta = -.17$ vs $-.21$).

In those who remained unimpaired at follow up (n = 146, mean interval = 383 days; SD = 52), test-retest reliability of the factor score was 0.86 (95% CI: .81, .90). Supplemental figure 8 provides information on test-retest reliability for all TabCAT-EXAMINER subtests. Annualized change scores were calculated for all participants. TabCAT-EXAMINER scores tended to decline more rapidly with more advanced disease progression (Figure 4). Examining the between-group differences, the asymptomatic group demonstrated significantly less annualized change than the prodromal group ($\beta = -.16$; 95% CI: $-.29$, $-.03$; $p = .02$) and the symptomatic group ($\beta = -.22$; 95% CI: $-.32$, $-.15$; $p < .001$). The comparison between prodromal and symptomatic groups was in the expected direction but the effect was small and not statistically significant ($\beta = -.11$; 95% CI: $-.29$, $-.07$; $p = .23$).

Asymptomatic participants (CDR®+NACC-FTLD = 0) who progressed to show at least mild symptoms(CDR®+NACC-FTLD≥0.5) at follow-up ($n = 14$) showed significantly greater annualized decline (moderate effect size) in TabCAT-EXAMINER scores ($\beta = -.34$; 95% CI: $-.52$, $-.16$; $p < .001$) than those who were unimpaired at both time points.

Discussion

We present the development and validation of the TabCAT-EXAMINER score. Fit statistics indicated the model fit the data well, with moderate-to-strong factor loadings for all subtests, and the normally distributed factor score showed high test-retest reliability. Our data support the construct validity and psychometrics of the TabCAT-EXAMINER score through strong

https://doi.org/10.1017/S1355617725000177 Published online by Cambridge University Press

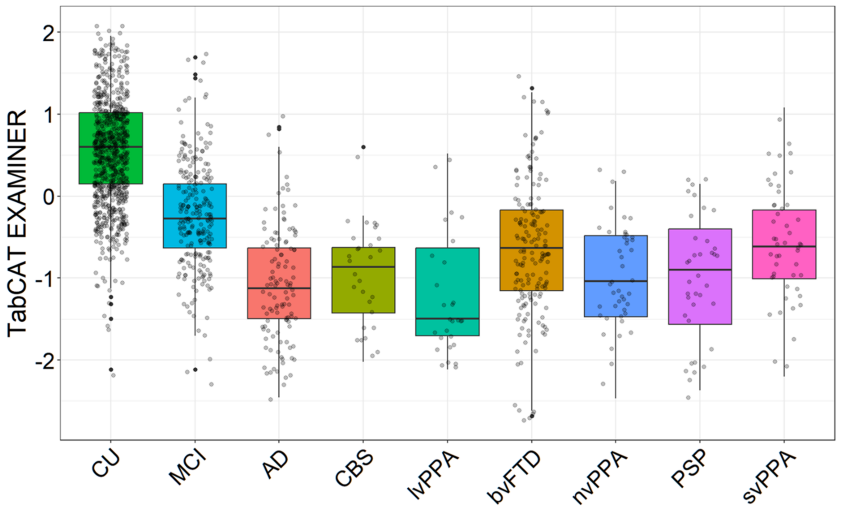


Figure 3. TabCAT-EXAMINER performance by clinical syndrome within the developmental sample. CU = cognitively unimpaired, MCI = mild cognitive impairment, lvPPA = logopenic primary progressive aphasia, AD = Alzheimer's disease dementia, CBS = corticobasal syndrome, PSP = progressive supranuclear palsy, nvPPA = nonfluent-variant primary progressive aphasia, svPPA = semantic-variant primary progressive aphasia, bvFTD = behavioral-variant frontotemporal dementia.

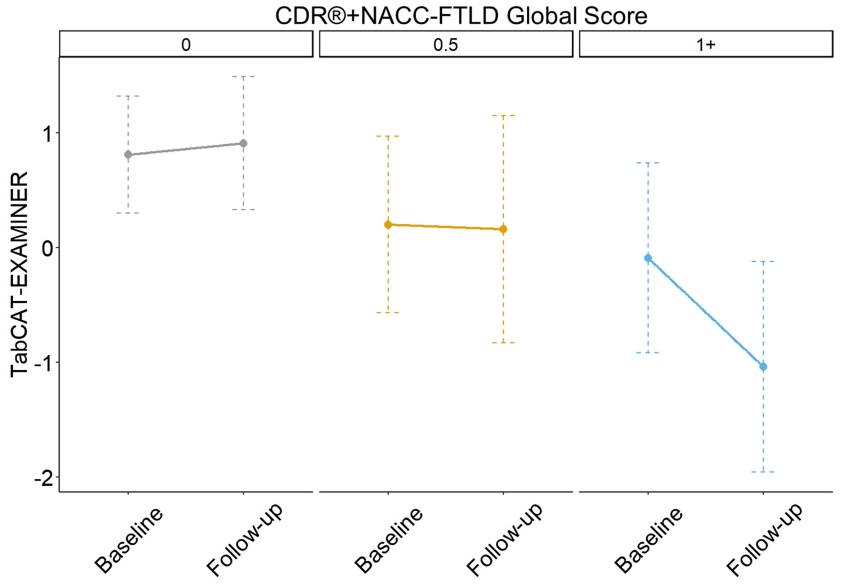


Figure 4. Longitudinal changes in TabCAT-EXAMINER scores by baseline disease severity. Presents the average TabCAT-EXAMINER score at baseline and follow-up by baseline CDR®+NACC-FTLD Global score (0 vs. 0.5 vs. 1+). Error bars represent standard deviation at each timepoint. Scores are in raw units.

associations with other tasks of executive functioning and processing speed. We also show known-group validity in that lower scores differentiated controls from multiple diagnostic syndromes and showed a stepwise decline with increasing disease severity. These results were replicated in a subsample of ALLFTD participants. Taken together, these data suggest the single and pragmatic TabCAT-EXAMINER score is a robust and valid measure of executive functioning across a range of neurodegenerative syndromes.

Several steps were taken to maximize the relevance of the TabCAT-EXAMINER score to a broad research community interested in studying executive functioning and processing speed. To improve the generalizability of the factor score for use across a range of clinical conditions, the development of the score was conducted within a diagnostically heterogeneous sample (i.e., disease severity, syndrome). Indeed, our results underscore the sensitivity of the TabCAT-EXAMINER score to executive deficits across a wide range of clinical syndromes. Additionally, the factor score was computed using IRT methodology to enhance its use by a variety of researchers with different goals. An important value of IRT is that valid and reliable scores on the same scale can be created using IRT regardless of which tests were administered, enabling researchers to tailor their battery to include only subsets that are pertinent to their study. The TabCAT-EXAMINER demonstrated this feature of IRT through several supplemental analyses that studied factor scores calculated in hypothetical scenarios in which only a subset of the tasks were administered. First, factor scores calculated with the full set of TabCAT-EXAMINER tasks were highly correlated with scores calculated with only a subset of test items. The variance around that association, however, shrank as more tests were included. Second, we investigated how the inclusion of specific items affected the association of the TabCAT-EXAMINER and the CDR®+NACC-FTLD Sum of Boxes. Once five of nine tasks were used to estimate the composite score, all generated scores showed similar associations with the CDR®+NACC-FTLD Sum of Boxes. Finally, the TabCAT-EXAMINER demonstrated strong test-retest reliability across time. Taken together, these analyses support use of the TabCAT-EXAMINER, even if data for all the subtasks are not available.

https://doi.org/10.1017/S1355617725000177 Published online by Cambridge University Press

We also conducted a series of analyses to validate the utility of these tests for FTD participants due to the importance of executive functioning within these syndromes. We first replicated prior investigations with the laptop-based NIH-EXAMINER, finding that the composite score captured executive functioning deficits of FTD, was sensitive to early stages of the disease, and was inversely related to disease severity as indexed by the negative correlation with the CDR®+NACC-FTLD, a measure developed specifically to assess clinical functioning in FTD. In a series of longitudinal analyses, faster rates of decline were observed in those with more advanced clinical severity and participants that showed clinical progression over 6 to 24 months exhibited greater declines in TabCAT-EXAMINER scores than those who remained clinically stable. Together, these results support this score as a tool for monitoring disease progression. Across both developmental and validation cohorts, asymptomatic f-FTLD mutation carriers performed worse on average than non-carrier family controls. Although these differences were non-significant and had small effect sizes, that is somewhat expected due to the small number of participants and the heterogeneity of this sample, which comprised three different pathogenic mutations that can present with phenotypic variability. Further study in larger samples is warranted to understand the potential of this score for identifying the earliest cognitive deficits in f-FTLD, which may vary by underlying pathogenic risk. The findings from this study of the TabCAT-EXAMINER score adds further support that digital cognitive tasks are reliable and may serve as effective outcomes for clinical trials (Inan et al., 2020; Staffaroni, Bajorek, et al., 2020; Staffaroni et al., 2024; Staffaroni, Tsoy, et al., 2020).

The code to generate these factor scores was written in an open-source software, *R*, allowing interested users to calculate factor scores using their own data. Additionally, the TabCAT platform is implementing an automatic calculation of the TabCAT-EXAMINER for users. Details about accessing TabCAT software can be found online at: https://tabcathealth.com/.

Our study has some limitations. Although the sample was diagnostically diverse, all participants completed testing in English, and participants were generally highly educated and predominantly white and non-Hispanic. Future studies with non-English speaking and culturally diverse samples are warranted to determine the generalizability of the current findings to other populations. Toward this end, an ongoing study is examining the TabCAT-EXAMINER score within a Latin American Cohort (ReDLat: Multi-Partner Consortium to Expand Dementia Research in Latin America; Ibanez et al., 2021). Our validation sample primarily comprised FTD participants; additional validation in other neurodegenerative and neurological populations would further demonstrate utility of the TabCAT-EXAMINER. The TabCAT-EXAMINER is not directly comparable to the original NIH-EXAMINER due to adaption changes (i.e., computer to tablet) and differences in specific components (e.g., Running dots, Match). A future cross-walk study will develop a method of converting between the executive composites.

Conclusion

Scores on the TabCAT-EXAMINER, a tablet-based battery of executive functioning tests, can be distilled into a single reliable metric with evidence of construct validity and utility for disease monitoring. Since executive functioning is impacted by a range of neurological and medical conditions, this score has broad relevance. These results add to an emerging body of literature supporting the use of digital health technologies for collecting clinical outcome assessments in AD and related dementia research. Future work is needed to understand the relevance of these tests to clinical care and validate their use in different languages and in underrepresented groups.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1355617725000177.

Funding statement. This work was supported by the National Institutes of Health [grants AG063911, AG62677, AG045390, NS092089, AG032306, AG016976, AG058233, AG038791, AG02350, AG019724, AG062422, NS050915, AG032289-11, AG077557, K23AG061253, K24AG045333, and U01NS128913], the Association for Frontotemporal Degeneration, the Bluefield Project to Cure FTD, the Rainwater Charitable Foundation, the Global Brain Health Institute, the Larry L. Hillblom Foundation [2014-A-004-NET], the John Douglas French Alzheimer's Foundation, and AlzOut. Samples from the National Centralized Repository for Alzheimer's Disease and Related Dementias, which receives government support under a cooperative agreement grant (U24 AG21886) awarded by the NIA, were used in this study. Dr Saloner. is supported by the Alzheimer's Association (AARF-23-1145318), New Vision Research (CCAD2024-001-1), American Academy of Neurology, Association for Frontotemporal Degeneration, and American Brain Foundation. Dr Boxer is supported by U19AG063911, R01AG073482, R01AG038791, and R01AG071756. Dr Kramer is supported by NIH (R01(s) AG032289 and AG048234) and Larry L. Hillblom Network Grant (2014-A-004-NET). We would also like to thank participants and collaborators from the ALLFTD (https://www.allftd.org/), which was funded by the National Institute on Aging and the National Institute of Neurological Diseases and Stroke (U19: AG063911).

Competing interests. Dr Staffaroni has provided consultation to Alector, Aviado Bio, CervoMed, Eli Lilly & Prevail Therapeutics, Passage Bio, and Takeda. He serves on the scientific review board for the Alzheimer's Drug Discovery Foundation. He is a co-inventor of four smartphone cognitive tasks (not included in the present study) and receives licensing fees. UCSF receives licensing fees for TabCAT that are used to maintain the software and its use. Dr Rojas is a site Principal Investigator for clinical trials sponsored by Eli-Lilly, Eisai and Amylyx. Dr Boxer has served as a consultant for Aeovian, AGTC, Alector, Arkuda, Arvinas, Boehringer Ingelheim, Denali, GSK, Life Edit, Humana, Oligomerix, Oscotec, Roche, TrueBinding, Wave, Merck and received research support from Biogen, Eisai and Regeneron.

References

Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology review*, *16*(1), 17–42.

Arrotta, K., Reyes, A., Kaestner, E., McDonald, C. R., Hermann, B. P., Barr, W. B., Sarmey, N., Sundar, S., Kondylis, E., Najm, I., Bingaman, W., & Busch, R. M. (2022). Cognitive phenotypes in frontal lobe epilepsy. *Epilepsia*, *63*(7), 1671–1681.

Assogna, M., Sprugnoli, G., Press, D., Dickerson, B., Macone, J., Bonnì, S., Borghi, I., Connor, A., Hoffman, M., Grover, N., Wong, B., Shen, C., Martorana, A., O'Reilly, M., Ruffini, G., El Fakhri, G., Koch, G., & Santarnecchi, E. (2021). Gamma-induction in frontotemporal dementia (GIFTeD) randomized placebo-controlled trial: Rationale, noninvasive brain stimulation protocol, and study design. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, *7*(1), e12219.

Azouvi, P., Arnould, A., Dromer, E., & Vallat-Azouvi, C. (2017). Neuropsychology of traumatic brain injury: An expert overview. *Revue Neurologique*, *173*(7-8), 461–472.

Barker, M. S., Gottesman, R. T., Manoochehri, M., Chapman, S., Appleby, B. S., Brushaber, D., Devick, K. L., Dickerson, B. C., Domoto-Reilly, K., Fields, J. A., Forsberg, L. K., Galasko, D. R., Ghoshal, N., Goldman, J., Graff-Radford, N. R., Grossman, M., Heuer, H. W., Hsiung, G.-Y., Knopman, D. S. … ALLFTD Consortium (2022). Proposed research criteria for prodromal behavioural variant frontotemporal dementia. *Brain*, *145*(3), 1079–1097.

Baudouin, A., Clarys, D., Vanneste, S., & Isingrini, M. (2009). Executive functioning and processing speed in age-related differences in memory: Contribution of a coding task. *Brain and cognition*, *71*(3), 240–245.

Bettcher, B. M., Mungas, D., Patel, N., Elofson, J., Dutt, S., Wynn, M., Watson, C. L., Stephens, M., Walsh, C. M., & Kramer, J. H. (2016). Neuroanatomical substrates of executive functions: Beyond prefrontal structures. *Neuropsychologia*, *85*, 100–109.

Boeve, B. F., Boxer, A. L., Rosen, H. J., Forsberg, L. K., Heuer, H. W., Brushaber, D., Appleby, B., Biernacka, J. M., Bordelon, Y. M., Botha, H., Brannelly, P., Dickerson, B. C., Dickson, D. W., Kimiko, D-Reilly, Faber, K., Fagan, A., Fields, J. A., Fishman, A., Foroud, T. M., . . . Wszolek, Z. (2020). Studying the natural history of frontotemporal lobar degeneration (FTLD): The ARTFL LEFFTDS longitudinal FTLD (ALLFTD) protocol: Neuropsychiatry and behavioral neurology: DLB and FTD—clinical manifestations. *Alzheimer's & Dementia*, *16*(S6), e045482.

Bott, N. T., Johnson, E. T., Schuff, N., Galifianakis, N., Subas, T., Pollock, J., Pressman, P., Kramer, J. H., & Possin, K. L. (2014). Sensitive measures of executive dysfunction in non-demented Parkinson's disease. *Parkinsonism & Related Disorders*, *20*(12), 1430–1433.

Brearly, T. W., Rowland, J. A., Martindale, S. L., Shura, R. D., Curry, D., & Taber, K. H. (2018). Comparability of iPad and web-based NIH Toolbox Cognitive Battery administration in veterans. *Archives of Clinical Neuropsychology*, *34*(4), 524–530.

Broadway, J. M., Rieger, R. E., Campbell, R. A., Quinn, D. K., Mayer, A. R., Yeo, R. A., Wilson, J. K., Gill, D., Fratzke, V., & Cavanagh, J. F. (2019). Executive function predictors of delayed memory deficits after mild traumatic brain injury. *Cortex*, *120*, 240–248.

Brooks, B. R., Miller, R. G., Swash, M., & Munsat, T. L. (2000). El Escorial revisited: Revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotrophic lateral sclerosis and other motor neuron disorders*, *1*(5), 293–299.

Byrne, B. M. (2001). Structural equation modeling: Perspectives on the present and the future. *International Journal of Testing*, *1*(3-4), 327–334.

Canini, M., Battista, P., Della Rosa, P. A., Catricalà, E., Salvatore, C., Gilardi, M. C., & Castiglioni, I. (2014). Computerized neuropsychological assessment in aging: Testing efficacy and clinical ecology of different interfaces. *Computational and Mathematical Methods in Medicine*, *2014*(1), 804723–13.

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of statistical Software*, *48*(6), 1–29.

Clark, D. J., Chatterjee, S. A., Skinner, J. W., Lysne, P. E., Sumonthee, C., Wu, S. S., Cohen, R. A., Rose, D. K., & Woods, A. J. (2021). Combining frontal transcranial direct current stimulation with walking rehabilitation to enhance mobility and executive function: A pilot clinical trial. *Neuromodulation: Technology at the Neural Interface*, *24*(5), 950–959.

Craft, S., Newcomer, J., Kanne, S., Dagogo-Jack, S., Cryer, P., Sheline, Y., Luby, J., Dagogo-Jack, A., & Alderson, A. (1996). Memory improvement following induced hyperinsulinemia in Alzheimer's disease. *Neurobiology of Aging*, *17*(1), 123–130.

Delis, D., Kramer, J., Kaplan, E., & Ober, B. (2000). *California verbal learning test, adult version (CVLT-II)*. The Psychological Corporation.

Dubois, B., Feldman, H. H., Jacova, C., DeKosky, S. T., Barberger-Gateau, P., Cummings, J., Delacourte, Aé, Galasko, D., Gauthier, S., Jicha, G., Meguro, K., O'Brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Stern, Y., Visser, P. J., & Scheltens, P. (2007). Research criteria for the diagnosis of Alzheimer's disease: Revising the NINCDS-ADRDA criteria. *The Lancet Neurology*, *6*(8), 734–746.

Faria, Cde A., Alves, H. V. D., & Charchat-Fichman, H. (2015). The most frequently used tests for assessing executive functions in aging. *Dementia & Neuropsychologia*, *9*(2), 149–155.

Foong, J., Rozewicz, L., Quaghebeur, G., Davie, C., Kartsounis, L., Thompson, A., Miller, H. D., & Ron, A. M. (1997). Executive function in multiple sclerosis. The role of frontal lobe pathology. *Brain: a journal of neurology*, *120*(1), 15–26.

Germine, L., Reinecke, K., & Chaytor, N. S. (2019). Digital neuropsychology: Challenges and opportunities at the intersection of science and software. *The Clinical neuropsychologist*, *33*(2), 271–286.

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, J. M., Rohrer, J. D., Black, S., Boeve, B. F., Manes, F., Dronkers, N. F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B. L., Knopman, D. S., Hodges, J. R., Mesulam, M. M., & Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*(11), 1006–1014.

Guarino, A., Favieri, F., Boncompagni, I., Agostini, F., Cantone, M., & Casagrande, M. (2019). Executive functions in Alzheimer disease: A systematic review. *Frontiers in Aging Neuroscience*, *10*, 437.

Hajjar, I., Okafor, M., McDaniel, D., Obideen, M., Dee, E., Shokouhi, M., Quyyumi, A. A., Levey, A., & Goldstein, F. (2020). Effects of candesartan vs lisinopril on neurocognitive function in older adults with executive mild cognitive impairment: A randomized clinical trial. *JAMA Network Open*, *3*(8), e2012252–e2012252.

Ibanez, A., Yokoyama, J. S., Possin, K. L., Matallana, D., Lopera, F., Nitrini, R., Takada, L. T., Custodio, N., Sosa Ortiz, A. L., Avila-Funes, Jé A., Behrens, M. I., Slachevsky, A., Myers, R. M., Cochran, J. N., Brusco, L. I., Bruno, M. A., Brucki, S. M. D., Pina-Escudero, S. D., Okada de Oliveira, M. . . . Miller, B. L. (2021). The multi-partner consortium to expand dementia research in Latin America (ReDLat): Driving multicentric research and implementation science. *Frontiers in Neurology*, *12*, 631722.

Inan, O. T., Tenaerts, P., Prindiville, S. A., Reynolds, H. R., Dizon, D. S., Cooper-Arnold, K., Turakhia, M., Pletcher, M. J., Preston, K. L., Krumholz, H. M., Marlin, B. M., Mandl, K. D., Klasnja, P., Spring, B., Iturriaga, E., Campo, R., Desvigne-Nickens, P., Rosenberg, Y., Steinhubl, S. R., & Califf, R. M. (2020). Digitizing clinical trials. *NPJ digital medicine*, *3*(1), 101.

Jack C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J. L., Montine, T., Phelps, C., Rankin, K. P., Rowe, C. C., Scheltens, P., Siemers, E., Snyder, H. M. . . . Silverberg, N. (2018). NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, *14*(4), 535–562.

Jaeger, J. (2018). Digit symbol substitution test: The case for sensitivity over specificity in neuropsychological testing. *Journal of Clinical Psychopharmacology*, *38*(5), 513–519.

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological bulletin*, *144*(11), 1147–1185.

Knopman, D. S., Kramer, J. H., Boeve, B. F., Caselli, R. J., Graff-Radford, N. R., Mendez, M. F., Miller, B. L., & Mercaldo, N. (2008). Development of methodology for conducting clinical trials in frontotemporal lobar degeneration. *Brain*, *131*(11), 2957–2968.

Kramer, J. H. (2014). Special series introduction: NIH EXAMINER and the assessment of executive functioning. *Journal of the International Neuropsychological Society*, *20*(1), 8–10.

Kramer, J. H., Mungas, D., Possin, K. L., Rankin, K. P., Boxer, A. L., Rosen, H. J., Bostrom, A., Sinha, L., Berhel, A., & Widmeyer, M. (2014). NIH EXAMINER: Conceptualization and development of an executive function battery. *Journal of the International Neuropsychological Society*, *20*(1), 11–19.

Leite, J., Gonçalves, Ó.F., & Carvalho, S. (2022). Speed of Processing (SoP) training plus α-tACS in people with mild cognitive impairment: A double blind, parallel, placebo controlled trial study protocol. *Frontiers in Aging Neuroscience*, *14*, 880510.

Lezak, M. D., Howieson, D. B., Loring, D. W., & Fischer, J. S. (2004). *Neuropsychological assessment*. Oxford University Press.

Litvan, I., MacIntyre, A., Goetz, C. G., Wenning, G. K., Jellinger, K., Verny, M., Bartko, J. J., Jankovic, J., McKee, A., Brandel, J. P., Chaudhuri, K. R., Lai, E. C., D'Olhaberriague, L., Pearce, R. K. B., & Agid, Y. (1998). Accuracy of the clinical diagnoses of Lewy body disease, Parkinson disease, and dementia with Lewy bodies: A clinicopathologic study. *Archives of neurology*, *55*(7), 969–978.

Longo, C. A., Kerr, E. N., & Smith, M. L. (2013). Executive functioning in children with intractable frontal lobe or temporal lobe epilepsy. *Epilepsy & Behavior*, *26*(1), 102–108.

Matsunaga, M. (2010). How to factor-analyze your data right: Do's, Don'ts, and how-to's. *International journal of psychological research*, *3*(1), 97–110.

https://doi.org/10.1017/S1355617725000177 Published online by Cambridge University Press

McKeith, I. G., Boeve, B. F., Dickson, D. W., Halliday, G., Taylor, J.-P., Weintraub, D., Aarsland, D., Galvin, J., Attems, J., Ballard, C. G., Bayston, A., Beach, T. G., Blanc, F., Bohnen, N., Bonanni, L., Bras, J., Brundin, P., Burn, D., Chen-Plotkin, A., … Kosaka, K. (2017). Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. *Neurology*, *89*(1), 88–100.

Miller, J. B., Banks, S. J., Léger, G. C., & Cummings, J. L. (2014). Randomized controlled trials in frontotemporal dementia: Cognitive and behavioral outcomes. *Translational Neurodegeneration*, *3*(1), 1–9.

Miyagawa, T., Brushaber, D., Syrjanen, J., Kremers, W., Fields, J., Forsberg, L. K., Heuer, W., H., Knopman, D., Kornak, J., Boxer, A., ARTFL/LEFFTDS Consortium. (2020). Use of the CDR® plus NACC FTLD in mild FTLD: data from the ARTFL/LEFFTDS consortium. *Alzheimer's & Dementia*, *16*, 79–90.

Miyagawa, T., Brushaber, D., Syrjanen, J., Kremers, W., Fields, J., Forsberg, L. K., Heuer, H. W., Knopman, D., Kornak, J., Boxer, A., Rosen, H. J., Boeve, B. F., Appleby, B., Bordelon, Y., Bove, J., Brannelly, P., Caso, C., Coppola, G., Dever, R., … Wszolek, Z. (2020). Utility of the global CDR® plus NACC FTLD rating and development of scoring rules: Data from the ARTFL/LEFFTDS Consortium. *Alzheimer's & Dementia*, *16*(1), 106–117.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex, frontal lobe, tasks: A latent variable analysis. *Cognitive psychology*, *41*(1), 49–100.

Monsell, S. E., Dodge, H. H., Zhou, X.-H., Bu, Y., Besser, L. M., Mock, C., Hawes, S. E., Kukull, W. A., & Weintraub, S. (2016). Results from the NACC uniform data set neuropsychological battery crosswalk study. *Alzheimer Disease & Associated Disorders*, *30*(2), 134–139.

Morris, J. C. (1997). Clinical dementia rating: A reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *International Psychogeriatrics*, *9*(S1), 173–176.

Mullady, S. S..-S., Castellanos, S., Lopez, L., Aguirre, G., Weeks, J., King, S., Valle, K., Goode, C., Tsoy, E., Possin, K., Miller, B., Kushel, M., & Lanata, S. (2022). Neurocognitive health of older adults experiencing homelessness in Oakland, California. *Frontiers in Neurology*, *13*, 905779.

Mungas, D., Reed, B. R., & Kramer, J. H. (2003). Psychometrically matched measures of global cognition, memory, and executive function for assesment of cognitive decline in older persons. *Neuropsychology*, *17*(3), 380–392.

Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S., Freedman, M., Kertesz, A., Robert, P. H., Albert, M., Boone, K., Miller, B. L., Cummings, J., & Benson, D. F. (1998). Frontotemporal lobar degeneration: A consensus on clinical diagnostic criteria. *Neurology*, *51*(6), 1546–1554.

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(2), 241–268.

Petersen, R. C., Aisen, P., Boeve, B. F., Geda, Y. E., Ivnik, R. J., Knopman, D. S., Mielke, M., Pankratz, V. S., Roberts, R., Rocca, W. A., Weigand, S., Weiner, M., Wiste, H., & Jack C. R. (2013). Mild cognitive impairment due to Alzheimer disease in the community. *Annals of Neurology*, *74*(2), 199–208.

Petersen, R. C., Aisen, P. S., Beckett, L. A., Donohue, M. C., Gamst, A. C., Harvey, D. J., C. R. Jack, Jagust, W. J., Shaw, L. M., Toga, A. W., Trojanowski, J. Q., & Weiner, M. W. (2010). Alzheimer's disease neuroimaging initiative (ADNI): Clinical characterization. *Neurology*, *74*(3), 201–209.

Petersen, R. C., Caracciolo, B., Brayne, C., Gauthier, S., Jelic, V., & Fratiglioni, L. (2014). Mild cognitive impairment: A concept in evolution. *Journal of Internal Medicine*, *275*(3), 214–228.

Piccinin, A. M., & Rabbitt, P. (1999). Contribution of cognitive abilities to performance and improvement on a substitution coding task. *Psychology and Aging*, *14*(4), 539–551.

Possin, K. L., Laluz, V. R., Alcantar, O. Z., Miller, B. L., & Kramer, J. H. (2011). Distinct neuroanatomical substrates and cognitive mechanisms of figure copy performance in Alzheimer's disease and behavioral variant frontotemporal dementia. *Neuropsychologia*, *49*(1), 43–48.

Possin, K. L., LaMarre, A. K., Wood, K. A., Mungas, D. M., & Kramer, J. H. (2014). Ecological validity and neuroanatomical correlates of the NIH EXAMINER executive composite score. *Journal of the International Neuropsychological Society*, *20*(1), 20–28.

Possin, K. L., Moskowitz, T., Erlhoff, S. J., Rogers, K. M., Johnson, E. T., Steele, N. Z., Higgins, J. J., Stiver J., Alioto A. G., Farias, S. T., Miller B. J., & Rankin, K. P. (2018). The brain health assessment for detecting and diagnosing neurocognitive disorders. *Journal of the American Geriatrics Society*, *66*(1), 150–156.

Possin, K. L., Moskowitz, T., Erlhoff, S. J., Rogers, K. M., Johnson, E. T., Steele, N. Z. R., Higgins, J. J., Stiver, J., Alioto, A. G., Farias, S. T., Miller, B. L., & Rankin, K. P. (2018). The brain health assessment for detecting and diagnosing neurocognitive disorders. *Journal of the American Geriatrics Society*, *66*(1), 150–156.

Postuma, R. B., Berg, D., Stern, M., Poewe, W., Olanow, C. W., Oertel, W., Obeso, Jé, Marek, K., Litvan, I., Lang, A. E., Halliday, G., Goetz, C. G., Gasser, T., Dubois, B., Chan, P., Bloem, B. R., Adler, C. H., & Deuschl, G. (2015). MDS clinical diagnostic criteria for Parkinson's disease. *Movement Disorders*, *30*(12), 1591–1601.

Rabinovici, G. D., Stephens, M. L., & Possin, K. L. (2015). Executive dysfunction. *CONTINUUM: Lifelong Learning in Neurology, 17(3): 423–433 Behavioral Neurology and Neuropsychiatry)*, 646.

Ramos, E., Dokuru, D., Van Berlo, V., Wojta, K., Wang, Q., & Huang, A. (2020). Genetic screening of a large series of North American sporadic and familial frontotemporal dementia cases. *Alzheimers Dement*, *16*(1), 118–130.

Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., van Swieten, J. C., Seelaar, H., Dopper, E. G. P., Onyike, C. U., Hillis, A. E., Josephs, K. A., Boeve, B. F., Kertesz, A., Seeley, W. W., Rankin, K. P., Johnson, J. K., Gorno-Tempini, M.-L., Rosen, H., … Miller, B. L. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, *134*(9), 2456–2477.

Reineberg, A. E., & Banich, M. T. (2016). Functional connectivity at rest is sensitive to individual differences in executive function: A network analysis. *Human Brain Mapping*, *37*(8), 2959–2975.

Roman, C. A., & Arnett, P. A. (2016). Structural brain indices and executive functioning in multiple sclerosis: A review. *Journal of Clinical and Experimental Neuropsychology*, *38*(3), 261–274.

Sachdev, P., Kalaria, R., O'Brien, J., Skoog, I., Alladi, S., Black, S. E., Blacker, D., Blazer, D. G., Chen, C., Chui, H., Ganguli, M., Jellinger, K., Jeste, D. V., Pasquier, F., Paulsen, J., Prins, N., Rockwood, K., Roman, G., & Scheltens, P. (2014). Diagnostic criteria for vascular cognitive disorders: A VASCOG statement. *Alzheimer Disease & Associated Disorders*, *28*(3), 206–218.

Salthouse, T. A. (1992). What do adult age differences in the digit symbol substitution test reflect? *Journal of Gerontology*, *47*(3), P121–P128.

Samejima, F. (2011). The general graded response model. In *Handbook of polytomous item response theory models* (pp. 87–118). Routledge.

Savalei, V., & Bentler, P. M. (2006). Structural equation modeling. In *The handbook of marketing research: Uses, misuses, and future advances*, vol. *330*, (pp. 36). SAGE Publications, Inc.

Schatz, J., Stancil, M., Katz, T., & Sanchez, C. E. (2014). EXAMINER executive function battery and neurologic morbidity in pediatric sickle cell disease. *Journal of the International Neuropsychological Society*, *20*(1), 29–40.

Schreiber, J. E., Possin, K. L., Girard, J. M., & Rey-Casserly, C. (2014). Executive function in children with attention deficit/hyperactivity disorder: The NIH EXAMINER battery. *Journal of the International Neuropsychological Society*, *20*(1), 41–51.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling*. Psychology press.

Staffaroni, A. M., Asken, B. M., Casaletto, K. B., Fonseca, C., You, M., Rosen, H. J., Boxer, A. L., Elahi, F. M., Kornak, J., Mungas, D., & Kramer, J. H. (2021). Development and validation of the Uniform Data Set (v3. 0) executive function composite score (UDS3-EF). *Alzheimer's & Dementia*, *17*(4), 574–583.

Staffaroni, A. M., Bajorek, L., Casaletto, K. B., Cobigo, Y., Goh, S.-Yang M., Wolf, A., Heuer, H. W., Elahi, F. M., Ljubenkov, P. A., Dever, R., Kornak, J., Appleby, B., Bove, J., Bordelon, Y., Brannelly, P., Brushaber, D., Caso, C., Coppola, G., Dheel, C., … ARTFL/LEFFTDS consortium (2020). Assessment of executive function declines in presymptomatic and mildly symptomatic familial frontotemporal dementia: NIH-EXAMINER as a potential clinical trial endpoint. *Alzheimer's & Dementia*, *16*(1), 11–21.

https://doi.org/10.1017/S1355617725000177 Published online by Cambridge University Press

Staffaroni, A. M., Clark, A. L., Taylor, J. C., Heuer, H. W., Sanderson-Cimino, M., Wise, A. B., Dhanam, S., Cobigo, Y., Wolf, A., Manoochehri, M., Forsberg, L., Mester, C., Rankin, K. P., Appleby, B. S., Bayram, E., Bozoki, A., Clark, D., Darby, R. R., Domoto-Reilly, K., … ALLFTD Consortium (2024). Reliability and validity of smartphone cognitive testing for frontotemporal lobar degeneration. *JAMA Network Open*, *7*(4), e244266–e244266.

Staffaroni, A. M., Quintana, M., Wendelberger, B., Heuer, H. W., Russell, L. L., Cobigo, Y., Wolf, A., Goh, S.-Y. M., Petrucelli, L., Gendron, T. F., Heller, C., Clark, A. L., Taylor, J. C., Wise, A., Ong, E., Forsberg, L., Brushaber, D., Rojas, J. C., VandeVrede, L., … Frontotemporal Dementia Prevention Initiative (FPI) Investigators (2022). Temporal order of clinical and biomarker changes in familial frontotemporal dementia. *Nature Medicine*, *28*(10), 2194–2206.

Staffaroni, A. M., Tsoy, E., Taylor, J., Boxer, A. L., & Possin, K. L. (2020). Digital cognitive assessments for dementia: Digital assessments may enhance the efficiency of evaluations in neurology and other clinics. *Practical Neurology (Fort Washington, Pa.)*, *2020*, 24–45.

Stasenko, A., Jacobs, D. M., Salmon, D. P., & Gollan, T. H. (2019). The Multilingual Naming Test (MINT) as a measure of picture naming ability in Alzheimer's disease. *Journal of the International Neuropsychological Society*, *25*(8), 821–833.

Stopford, C. L., Thompson, J. C., Neary, D., Richardson, A. M., & Snowden, J. S. (2012). Working memory, attention, and executive function in Alzheimer's disease and frontotemporal dementia. *Cortex*, *48*(4), 429–446.

Team, R. C. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Vaportzis, E., Giatsi Clausen, M., & Gow, A. J. (2017). Older adults perceptions of technology and barriers to interacting with tablet computers: A focus group study. *Frontiers in Psychology*, *8*, 1687.

Vossel, K., Ranasinghe, K. G., Beagle, A. J., La, A., Ah Pook, K., Castro, M., Mizuiri, D., Honma, S. M., Venkateswaran, N., Koestler, M., Zhang, W., Mucke, L., Howell, M. J., Possin, K. L., Kramer, J. H., Boxer, A. L., Miller, B. L., Nagarajan, S. S., & Kirsch, H. E. (2021). Effect of levetiracetam on cognition in patients with Alzheimer disease with and without epileptiform activity: A randomized clinical trial. *JAMA Neurology*, *78*(11), 1345–1354.

Weintraub, S., Besser, L., Dodge, H. H., Teylan, M., Ferris, S., Goldstein, F. C., Giordani, B., Kramer, J., Loewenstein, D., Marson, D., Mungas, D., Salmon, D., Welsh-Bohmer, K., Zhou, X.-H., Shirk, S. D., Atri, A., Kukull, W. A., Phelps, C., & Morris, J. C. (2018). Version 3 of the Alzheimer disease centers' neuropsychological test battery in the Uniform Data Set (UDS). *Alzheimer Disease & Associated Disorders*, *32*(1), 10–17.

Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., Carlozzi, N. E., Slotkin, J., Blitz, D., Wallner-Allen, K., Fox, N. A., Beaumont, J. L., Mungas, D., Nowinski, C. J., Richler, J., Deocampo, J. A., Anderson, J. E., Manly, J. J., Borosh, B., … Gershon, R. C. (2013). Cognition assessment using the NIH Toolbox. *Neurology*, *80*(11_supplement_3), S54–S64.

Weintraub, S., Salmon, D., Mercaldo, N., Ferris, S., Graff-Radford, N. R., Chui, H., Cummings, J., DeCarli, C., Foster, N. L., Galasko, D., Peskind, E., Dietrich, W., Beekly, D. L., Kukull, W. A., & Morris, J. C. (2009). The Alzheimer's Disease Centers' Uniform Data Set (UDS): The neuropsychological test battery. *Alzheimer Disease & Associated Disorders*, *23*(2), 91–101.

Williams, B., Onsman, A., & Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine*, *8*, 1–13.

Wilson, S. A., Byrne, P., Rodgers, S. E., & Maden, M. (2022). A systematic review of smartphone and tablet use by older adults with and without cognitive impairment. *Innovation in Aging*, *6*(2), igac002.

You, S. C., Geschwind, M. D., Sha, S. J., Apple, A., Satris, G., Wood, K. A., Johnson, E. T., Gooblar, J., Feuerstein, J. S., Finkbeiner, S., Kang, G. A., Miller, B. L., Hess, C. P., Kramer, J. H., & Possin, K. L. (2014). Executive functions in premanifest Huntington's disease. *Movement Disorders*, *29*(3), 405–409.

Yu, F., Lin, F. V., Salisbury, D. L., Shah, K. N., Chow, L., Vock, D., Nelson, N. W., Porsteinsson, A. P., & Jack C. (2018). Efficacy and mechanisms of combined aerobic exercise and cognitive training in mild cognitive impairment: Study protocol of the ACT trial. *Trials*, *19*(1), 1–13.

https://doi.org/10.1017/S1355617725000177 Published online by Cambridge University Press