



RESEARCH ARTICLE

# Feasibility, validity, and normative data for the remote Uniform Data Set neuropsychological battery at the University of Pennsylvania Alzheimer's Disease Research Center

Katherine Hackett<sup>1</sup>  | Yidan Shi<sup>2</sup> | Laura Schankel<sup>3</sup> | Nicole Oliveira<sup>3</sup> |  
Melissa Kelley<sup>3</sup> | Hannah McCoubrey<sup>3</sup> | Sara Manning Peskin<sup>3</sup> | Kyra O'Brien<sup>3</sup> |  
Sharon X. Xie<sup>2</sup> | David Wolk<sup>3</sup> | Dawn Mechanic-Hamilton<sup>3</sup> 

<sup>1</sup>Department of Psychology and Neuroscience, Temple University, Philadelphia, Pennsylvania, USA

<sup>2</sup>Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

<sup>3</sup>Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

## Correspondence

Dawn Mechanic-Hamilton, Department of Neurology, Perelman School of Medicine, University of Pennsylvania, 3400 Civic Center Blvd, Philadelphia, PA 19147, USA.

Email:

[dawn.mechanic@penmedicine.upenn.edu](mailto:dawn.mechanic@penmedicine.upenn.edu)

## Present address

Katherine Hackett, Division of General Internal Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

Yidan Shi, Division of Biostatistics, Department of Population Health, NYU Grossman School of Medicine, New York, NY 10016, USA

## Funding information

National Institutes of Health (NIH); National Institute on Aging, Grant/Award Numbers: F31AG069444, T32AG066598, K23AG065499, P30AG072979

## Abstract

**INTRODUCTION:** We investigated the feasibility and validity of the remotely-administered neuropsychological battery from the National Alzheimer's Coordinating Center Uniform Data Set (UDS T-Cog).

**METHODS:** Two hundred twenty Penn Alzheimer's Disease Research Center participants with unimpaired cognition, mild cognitive impairment, and dementia completed the T-Cog during their annual UDS evaluation. We assessed administration feasibility and diagnostic group differences cross-sectionally across telephone versus videoconference modalities, and compared T-Cog to prior in-person UDS scores longitudinally.

**RESULTS:** Administration time averaged 54 min and 79% of participants who initiated a T-Cog completed all 12 subtests; completion time and rates differed by diagnostic group but not by modality. Performance varied expectedly across groups with moderate to strong associations between most T-Cog measures and in-person correlates, although select subtests demonstrated lower comparability.

**DISCUSSION:** The T-Cog is feasibly administered and shows preliminary validity in a cognitively heterogeneous cohort. Normative data from this cohort should be expanded to more diverse populations to enhance utility and generalizability.

## KEYWORDS

accessibility, Alzheimer's Disease Research Center, COVID-19, digital neuropsychology, normative data, T-Cog, teleneuropsychology, Uniform Data Set

## Highlights

- This study examined the feasibility and validity of the remote Uniform Data Set (also known as the T-Cog) and contributes key normative data for widespread use.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* published by Wiley Periodicals, LLC on behalf of Alzheimer's Association.

- A remote neuropsychological battery was feasibly administered with high overall engagement and completion rates, adequate reliability compared to in-person testing, and evidence of validity across diagnostic groups.
- Typical barriers to administration included hearing impairment, technology issues, and distractions; hearing difficulties were particularly common among cognitively impaired groups.
- Certain tests were less closely related to their in-person correlates and should be used with caution.

## 1 | INTRODUCTION

In response to the coronavirus disease 2019 (COVID-19) pandemic, in-person clinical and research activities around the world came to a halt. As clinicians grappled with delivering quality health care while minimizing exposure risk, researchers of longitudinal trials were faced with the challenge of maintaining data collection.<sup>1</sup> It was unclear whether shifting to remote protocols would be (a) worthwhile given the uncertain duration of the pandemic, (b) feasible, and (c) valid. Limited normative data for remotely administered neuropsychological measures posed an additional challenge in characterizing clinical groups and maintaining the longitudinal fidelity of research databases.<sup>2</sup> The National Alzheimer's Coordinating Center (NACC) for the National Institute on Aging (NIA) Alzheimer's Disease Research Centers (ADRCs) responded to this dilemma. In June 2020, members of the NACC ADRC Clinical Task Force Cognitive Working Group worked to quickly expand telehealth data collection and shift administration of the standardized annual evaluation (the Uniform Data Set version 3; UDSv3) to remote formats. This included a remote neuropsychological battery, also known as the T-Cog, whose preliminary feasibility and test-retest reliability has been recently reported by Sachs et al. (2024) and Howard et al. (2023).<sup>3,4</sup> Herein, we expand upon the feasibility and validity findings surrounding use of the T-Cog in a longitudinal sample at the University of Pennsylvania (Penn ADRC).

### 1.1 | Brief background of teleneuropsychology

Although the COVID-19 pandemic was an impetus for increased telehealth utilization, digital platforms for neuropsychological assessment had emerged several years prior.<sup>5</sup> Teleneuropsychology, initially defined by Cullum and Grosch (2013) and later by Bilder and colleagues (2020), was seen as a means to increase access gaps due to geographical or mobility barriers.<sup>6-8</sup> A handful of prepandemic studies compared in-person to remote assessments and generally demonstrated adequate concurrent validity, including strong within-person, across-modality correlations for tests of global cognition, memory, attention, and language among healthy adults and those with varying degrees of cognitive impairment<sup>9-11</sup>; see Bready et al. (2017) and Marra et al. (2020) for systematic reviews.<sup>12,13</sup>

Teleneuropsychology expanded dramatically in response to the COVID-19 pandemic. According to a survey by Fox-Fuller and col-

leagues (2022) including 87 licensed and nonlicensed U.S. professionals using teleneuropsychology with adults, 82% of respondents had used teleneuropsychology only since the COVID-19 pandemic.<sup>2</sup> The predominant devices used were computers (84%), tablets (52%), and telephone audio calls (43%). Given the postpandemic uptick in remote assessment, the Inter Organizational Practice Committee released guidance including strategies to select appropriate measures, logistical considerations, and processes to weigh the risks and benefits of remote administration for each examinee.<sup>6,14</sup> Still, hesitations remain among neuropsychologists who are more confident in remote formats for history taking or feedback rather than standardized assessment.<sup>15,16</sup> Most published teleneuropsychology studies are limited to clinic-to-clinic settings, whereas very few have examined home-based remote assessment where noncontrolled environments can pose additional confounds.<sup>2,6,13,17-20</sup> Given the ethical concerns surrounding limited computer and internet access,<sup>13,21</sup> an additional gap in the teleneuropsychology literature includes telephone-based modalities that may be more accessible and familiar than videoconference platforms among older adults and individuals from lower socioeconomic backgrounds.<sup>1,6</sup>

Herein, we contribute insights from over 1-year of remote assessment of the UDS in a heterogeneous group of ADRC research participants across the cognitive spectrum, adding to the literature on the feasibility and validity of remote neuropsychological assessment of older adults. We provide normative data for use in other clinical and research settings with similar demographic characteristics. In addition to cross-sectional results from administration of the T-Cog at a single timepoint, we present longitudinal results from a subset of well-characterized clinical research participants to examine the concurrent validity and reliability of remote (T-Cog) versus prior in-person testing. Our findings add to the literature on remote neuropsychological assessment and help inform the appropriateness of continued use of the T-Cog within clinical and research settings.

## 2 | METHODS

### 2.1 | Study design and population

The Penn ADRC recruits participants from the Penn Memory Center and the Philadelphia community to join the Clinical Core cohort and complete annual visits. The UDS includes collection of medical and psychiatric history, neurologic exam, self- and partner-report of

## RESEARCH IN CONTEXT

- 1. Systematic review:** The authors of this study reviewed literature concerning teleneuropsychology, its evolution in response to the coronavirus disease 2019 (COVID-19) pandemic, and applications for Alzheimer's disease (AD) longitudinal trials.
- 2. Interpretation:** The remotely-administered Uniform Data Set (UDS; T-Cog) was feasibly administered over the phone or videoconference to older adult research participants across the cognitive aging spectrum and demonstrated evidence of preliminary validity. Common barriers to administration included hearing impairment, technical difficulties, and distractions. Feasibility was relatively lower among participants with dementia-level cognitive impairment. Select subtests, including Oral Trail Making Test, demonstrated lower feasibility, validity, and reliability.
- 3. Future directions:** Preliminary findings and normative data presented here should be expanded to include more diverse populations to enhance the accessibility and generalizability of remote neuropsychological services in AD research studies.

cognitive and neuropsychiatric symptoms, and a battery of neuropsychological tests.<sup>22</sup> Consensus diagnosis is completed on a yearly basis based on the above information by a panel of clinicians. In response to COVID-19, NACC compiled a remote version of the neuropsychological battery (T-Cog) using tests with previously demonstrated feasibility in remote administration and to allow for both telephone and video formats (see Figure 1).

Administration of the T-Cog began in June 2020 at the Penn ADRC. Although the T-Cog is still used in limited cases, most initial T-Cog administrations occurred between June 2020 and September 2021. All visits were conducted in English and only among participants who are fluent in and prefer English testing, as Spanish-speaking examiners were not available at that time. As of September 2021, more than 200 Penn ADRC participants had completed the T-Cog, and many had completed the in-person UDS in prior years. The present observational study focuses on individuals who met criteria for unimpaired cognition (UC), mild cognitive impairment (MCI), or dementia due to Alzheimer's Disease (AD) based on consensus diagnosis following completion of the T-Cog. We also examined data from a subset who completed prior annual in-person UDS testing approximately  $\approx 2$  years (T1) and 1-year (T2) before their remote T-Cog testing (T3) (see Figure 2; average duration between T3 and T2 (means  $\pm$  SD) =  $426 \pm 56$  days; T2 and T1 =  $381 \pm 50$  days). These longitudinal data spanning three consecutive timepoints allowed us to account for random within-person effects for all remaining analyses. Participants who reverted to UC (e.g., assigned a consensus diagnosis of MCI at T2 then UC at T3) were

excluded from the sample. Participants who progressed ( $n = 12$ ) were included and their diagnosis at the most recent timepoint (T3) was used.

## 2.2 | Administration of cognitive, psychological, and functional measures

The T-Cog includes measures spanning domains of global cognition (Montreal Cognitive Assessment—Blind (MoCA-B)),<sup>23,24</sup> verbal memory (Craft Story 21 Immediate and Delayed Recall [Story Recall Immediate/Delay]),<sup>25</sup> attention/working memory (Number Span Forward & Backward [Digits Forwards/Backwards]),<sup>22</sup> language (Verbal Naming Test; Letter and Category fluency),<sup>26,22</sup> and executive functioning and processing speed (Oral Trail Making Test Parts A/B [Oral Trails A/B]).<sup>27</sup> As shown in Figure 1, the T-Cog does not include measures of visuospatial skills or visual memory due to constraints of the administration format.

T-Cog examiners were Bachelors-level or Master's-level clinical research coordinators (CRCs) trained to administer the T-Cog according to standardized procedures and demonstrated proficiency through a certification process, consistent with training employed for the traditional in-person UDS. The T-Cog manual provided by NACC (naccdata.org) includes detailed instructions and prompts at the beginning of testing to ensure basic requirements for valid administration are met, including minimizing distractions, ensuring privacy, confirming adequate hearing and/or vision, and agreeing to complete the evaluation independently. Modality of testing (telephone vs videoconference) was determined in advance by asking participants to designate their preference and confirm access to requisite technology (e.g., device with or without video capability, internet access). If completing testing via videoconference, the study team sent a Health Insurance Portability and Accountability Act (HIPAA)-compliant video call link to the participant's established email address.

Neuropsychiatric symptoms were measured with the Geriatric Depression Scale (GDS; self-report)<sup>28</sup> and Neuropsychiatric Inventory (NPI-Q; partner report).<sup>29</sup> Everyday functioning was reported by the study partner on the Functional Rating Scale (FRS)<sup>30</sup> and the Functional Assessment Scale (FAS).<sup>31</sup> Global cognitive and functional abilities were also assessed via the Clinical Dementia Rating (CDR) scale.<sup>32</sup>

## 2.3 | Statistical analysis

### 2.3.1 | Feasibility: Is it feasible to administer the UDS remotely to individuals across the cognitive aging spectrum?

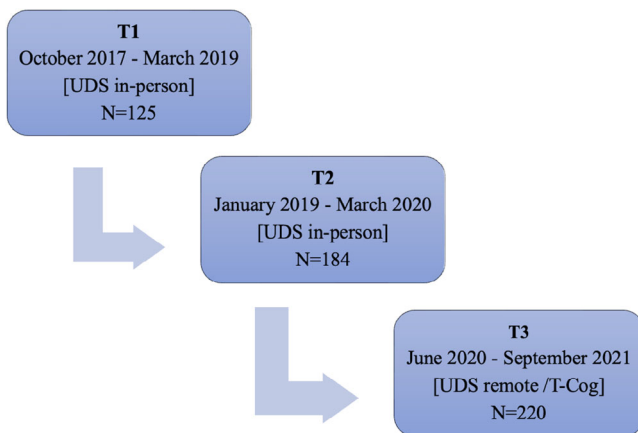
Rates of initiating versus declining to engage in remote testing were collected among participants who completed a portion of their annual UDS visit within the T3 study period. Among those initiating a

Domain	UDS v3 in-person	UDS v3 T-Cog
Global Cognition	MoCA	MoCA-Blind <sup>a</sup>
Verbal Memory	Craft Story 21 Immediate & Delayed Recall	Craft Story 21 Immediate & Delayed Recall
Visual Memory	Benson Complex Figure Delayed Recall & Recognition	
Attention & Working Memory	Number Span Forward & Backward	Number Span Forward & Backward
Visuospatial	Benson Complex Figure Copy MoCA Visuospatial items	
Language	Multilingual Naming Test (MINT)	Verbal Naming Test
Verbal Fluency	Letter (F, L) & Category (Animals, Vegetables) Fluency	Letter (F, L) & Category (Animals, Vegetables) Fluency
Executive Functioning & Processing Speed	Trail Making Test Parts A/ B	Oral Trail Making Test Parts A/B

<sup>a</sup>Excludes naming, visuospatial and set switching items from the MoCA

	Difference in modality only
	Difference in cognitive process
	Unavailable via remote administration

**FIGURE 1** Comparison of in-person versus remote UDS measures by cognitive domain. UDS, Uniform Data Set.



**FIGURE 2** Flowchart depicting repeat UDS evaluations including two in-person (T1 and T2) and one remote (T3; T-Cog) timepoint. The average duration between T3 and T2 was  $426 \pm 56$  days, and between T2 and T1 was  $381 \pm 50$  days.

T-Cog, administration data from individuals with diagnoses of UC, MCI, or AD were collected. Descriptive statistics were used to characterize feasibility of the administration, including duration, modality of administration (telephone vs videoconference), examiner-assigned validity ratings, and completion rates. To determine the impact of

diagnostic group and modality of administration on feasibility outcomes, results were examined across groups using between-groups analysis of variance (ANOVA), chi-square analyses, or nonparametric equivalents when appropriate. Contributors to poor test validity were probed by comparing valid versus invalid administrations across several participant features using ANOVA and chi-square analyses.

### 2.3.2 | Group differences: Does performance on the T-Cog vary across diagnostic groups as expected?

Performance on T-Cog measures at T3 was examined across the three diagnostic groups to test for expected decrements across increasing levels of cognitive impairment. Because variables were approximately normally distributed, univariate linear regression analyses were used, covarying for age, sex, race, education, and depression. These covariates were chosen a priori given their known associations with neuropsychological test performance.<sup>33-38</sup> Analyses were repeated with and without participants whose administration was rated as “questionably valid” or “invalid,” across all modalities, and stratified by telephone versus videoconference modality. We hypothesized that performance on the T-Cog would differ significantly across diagnostic groups, with the UC group performing the best and the AD group performing the worst.

### 2.3.3 | Concurrent validity and reliability in asymptomatic individuals: Is performance on the T-Cog comparable to performance on prior in-person testing among those with UC?

To examine concurrent validity, we compared scores on the T-Cog at T3 to the in-person UDS in the UC group, excluding questionably invalid or invalid T-Cog results. First, related-samples Wilcoxon signed-rank tests were used to compare features of the UC group across timepoints T2 and T3 to explore whether participants changed during this timeframe. Next, we examined Spearman's correlation coefficients between in-person (T2) and remote (T3) UDS tests and used two-way mixed effects, absolute agreement, single rater intraclass correlation coefficients (ICC)<sup>39</sup> to compare the reliability of in-person (T1) versus in-person (T2) performance to in-person (T2) versus remote (T3) performance. Given the expected weaker associations between tests that are not comparable in score range and underlying cognitive process (e.g., Oral vs Written Trail Making test), subsequent analyses were restricted to tests with an equivalent score range (green rows in Figure 1; MoCA-B to MoCA conversion was used for longitudinal analyses to facilitate direct comparison<sup>40</sup>). Repeated-measures ANOVA with covariates (age at T2, sex, race, education, and change in depression) were used to test for differences in UDS performance at T2 versus T3. Finally, linear mixed-effects models (LMMs) with restricted maximum likelihood (REML) estimation were used to examine the difference in UDS test scores at time T2 versus T3. The models include timepoint, modality, the interaction between timepoint and modality, and controlled covariates (age, sex, race, education, and depression) as fixed effects, and subject-specific random intercepts. We hypothesized that UCs would perform similarly on UDS tests at in-person versus remote formats, without a significant effect of timepoint.

### 2.3.4 | Concurrent validity in symptomatic individuals: Are rates of change on the UDS comparable across timepoints regardless of in-person versus remote administration?

Given the progressive nature of cognitive decline in symptomatic groups (MCI and AD), we expect a decline in cognitive test performance over time. Therefore, comparing UDS scores across only two timepoints does not allow us to determine equivalency in administration format among symptomatic groups. Instead, we investigated whether rates of change on UDS scores from T1 (in-person) to T2 (in-person) differed significantly from rates of change from T2 (in-person) to T3 (remote) within each diagnostic group. Consistent rates of change would indicate negligible differences due to remote administration format at T3. We hypothesized that the rate of decline from T2 → T3 may be greater than T1 → T2 among the MCI and AD groups, due either to increased difficulty of remote testing at T3 or to increased impact of the COVID-19 pandemic on symptomatic groups. UDS change scores were calculated by subtracting the raw score of each test at the earlier timepoint from the subsequent timepoint. Repeated-measures

ANOVA of T1 → T2 versus T2 → T3 change scores were used to test for differences in rates of change within each diagnostic group. If change scores were statistically significant, clinical significance was determined if change scores differed by  $\geq 0.5$  SD of the corresponding measure within that diagnostic group at baseline, which is an established form of minimally clinically important difference (MCID).<sup>41</sup> Covariates included age at T2, sex, race, education, and average depression. All statistical tests were two-sided. Statistical significance was set at the 0.05 level unless otherwise noted.

## 3 | RESULTS

A total of 220 participants completed the remote UDS T-Cog at T3, including  $n = 156$  UC,  $n = 26$  MCI, and  $n = 38$  AD. As shown in Table 1, participants ranged in age from 57 to 93 years of age ( $M = 74.2 \pm 6.5$ ), were highly educated ( $M = 16.6 \pm 2.7$  years), and a majority identified as female (58%), White (78%), and non-Hispanic/Latinx (98.2%). Given the small sample size of Asian ( $n = 1$ ) and Multiracial ( $n = 4$ ) participants, race was dichotomized into White and non-White categories to improve statistical power and interpretability in subsequent analyses. Between-group ANOVA revealed a significant difference between diagnostic groups for age ( $p = 0.012$ ), with post hoc comparisons revealing that the AD group was significantly older than the UC group. The chi-square analyses also revealed a significant difference in distribution of sex ( $p = 0.005$ ), with a larger proportion of female participants in the UC group compared to MCI and AD. Significant differences on other key nondemographic variables were observed in expected directions, such that more impairment on measures of global cognition, function, and mood were observed with increased diagnostic severity.

### 3.1 | Feasibility

Of all the participants completing a portion of their annual visit within the T3 study period (June 2020 to September 2021), 9% did not initiate the T-Cog due to appointment no shows/cancellations or a team decision that testing would be inappropriate due to level of impairment, whereas only 2% declined due to unwillingness to engage in remote testing. The T-Cog was most frequently administered via videoconferencing (50%) or telephone (48%), with a small number administered via combination of videoconferencing and telephone (2%). Age, diagnosis, global cognition (MoCA-B), function (CDR, FRS, FAS), and mood (GDS, NPI-Q) did not differ across modality of administration; however, those participants electing the telephone over videoconferencing modality had lower education, were more likely to be in the non-White and Hispanic/Latinx racial and ethnic categories, and were more likely to be female ( $p$ 's < 0.05). Across all modalities and all diagnostic groups, time to completion ranged from 30 to 120 min ( $M = 54.35 \pm 14.2$  min). Significant ( $p < 0.001$ ) differences were observed between diagnostic groups such that time to completion was highest among the AD group ( $64.4 \pm 20.6$  min), followed by MCI ( $56.9 \pm 11.5$  min) and UC ( $51.4 \pm 11.1$ ). Time to completion did not differ significantly across



**TABLE 1** Participant descriptives at T-Cog administration (T3) by diagnostic group—all modalities.

	UC (71%; n = 156)		MCI (12%; n = 26)		AD (17%; n = 38)		Total (N = 220)		p-value
	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	
Age	73.34 (5.6)	62–93	76.23 (8.1)	58–88	76.18 (8.2)	57–90	74.17 (6.5)	57–93	<b>0.012<sup>b</sup></b>
Education	16.54 (2.6)	9–22	16.77 (3.3)	8–20	16.58 (2.9)	9–20	16.57 (2.7)	8–22	0.924
Sex (%F)	65%		50%		37%		58%		<b>0.005</b>
Race (White/Black/ Asian/Multiracial)	73%/24%/ < 1%/3%		85%/15%/0%/0%		92%/8%/0%/0%		78%/20%/ < 1%/2%		0.254
Ethnicity (%Hispanic /Latinx)	2.6%		0%		0%		1.82%		0.434
MoCA-B	19.97 (1.8)	13–22	16.48 (2.7)	12–20	9.14 (4.1)	2–16	17.72 (4.7)	2–22	<b>&lt;0.001<sup>a,b,c</sup></b>
CDR Global	0.077 (.18)	0–0.5	0.519 (.10)	0.5–1.0	1.22 (.68)	0.5–3.0	.327 (.54)	0–3	<b>&lt;0.001<sup>a,b,c</sup></b>
GDS	1.17 (1.6)	0–10	2.25 (2.2)	0–7	2.41 (2.4)	0–7	1.51 (1.9)	0–10	<b>0.001<sup>a,b</sup></b>
FRS	1.20 (1.7)	0–9	7.88 (4.7)	0–21	19.71 (10.1)	2–40	5.90 (8.9)	0–40	<b>&lt;0.001<sup>a,b,c</sup></b>
FAS	0.17 (.53)	0–3	6.20 (5.4)	0–23	17.03 (7.8)	2–29	3.89 (7.4)	0–29	<b>&lt;0.001<sup>a,b,c</sup></b>
NPI-Q	0.28 (.88)	0–6	2.54 (3.0)	0–10	4.37 (2.9)	0–11	1.31 (2.4)	0–11	<b>&lt;0.001<sup>a,b,c</sup></b>
T-Cog time to completion	51.37 (11.1)	32–90	56.94 (11.5)	40–80	64.41 (20.6)	30–120	54.35 (14.2)	30–120	<b>&lt;0.001<sup>b</sup></b>
% Completing all 12 measures	92%		81%		26%		79%		<b>&lt;0.001</b>

Note: p-value according to between groups ANOVA or Pearson chi-square test, with Bonferroni-adjusted post hoc tests.

Bold values are statistically significant.

Abbreviations: AD, dementia due to Alzheimer's disease; ANOVA, analysis of variance; CDR, Clinical Dementia Rating; FAS, Functional Assessment Scale; FRS, Functional Rating Scale; GDS, Geriatric Depression Scale; UC, unimpaired cognition; MCI, mild cognitive impairment; MoCA-B, Montreal Cognitive Assessment-Blind; NPI-Q, Neuropsychiatric Inventory.

<sup>a</sup>UC versus MCI.

<sup>b</sup>UC versus AD.

<sup>c</sup>MCI versus AD.

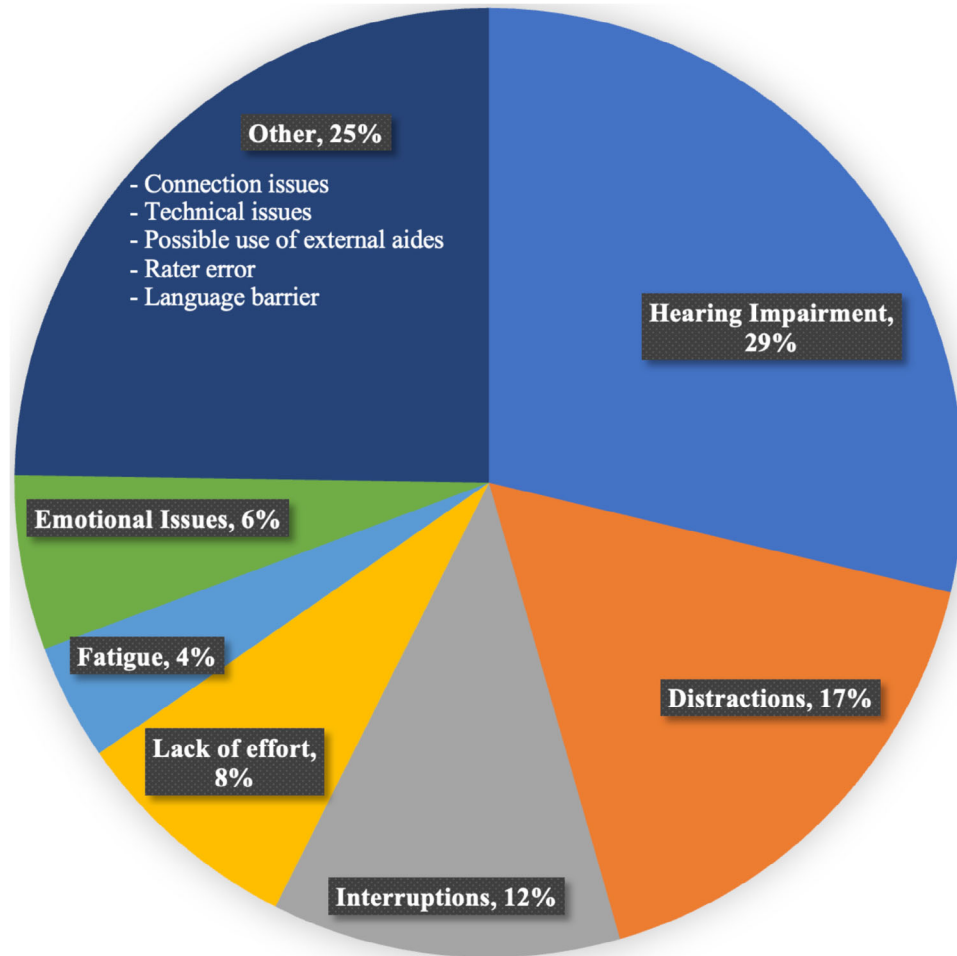
modality. Across the entire sample, 79% of the 220 participants initiating a T-Cog completed all 12 tests (Table 1); overall completion rates did not differ across modality. The proportion of participants completing all 12 measures was highest among the UC (92%) and MCI (81%) groups, followed by the AD group (26%; Table 1). Only 1% of the total sample (n = 3) completed less than half of the T-Cog tests; these individuals were all in the AD group and discontinued due to comprehension difficulties. Across all modalities and all diagnostic groups, all individual tests had >97% completion rate, with the exception of Oral Trails B, which had the lowest completion rate at 84%, followed by Verbal Naming Test at 93% (results were similar when stratified by modality, with slightly lower completion rates within the videoconference modality; see Table A1).

Examiners were trained to assign validity ratings to characterize the overall administration of each T-Cog according to the following options: (1) very valid, (2) questionably valid, or (3) invalid. Among the current sample, 82% of T-Cog administrations were rated as “very valid,” whereas a minority were rated as “questionably valid” (17%) and “invalid” (1%); the proportion of valid ratings did not differ across modality. Among the 18% rated as “questionably valid” or “invalid,” the most common contributors to poor validity included hearing impairment (29%), “other” issues including connectivity difficulties (25%), and distractions (17%; see Figure 3). The most common contributor to poor

validity among the MCI and AD groups specifically was hearing impairment (55%). Compared to individuals within the “questionably valid” or “invalid” groups, individuals in the “very valid” group demonstrated significantly higher MoCA-B scores, shorter time to completion, less functional impairment on the FRS and FAS, and included a higher proportion of UC participants (see Table A2). Participants in the “very valid” group did not differ on other demographic variables including age, sex, education, race, ethnicity, depression, neuropsychiatric symptoms, or mode of communication (telephone vs videoconference; p-values > 0.05).

### 3.2 | Diagnostic group differences

Performance on the T-Cog varied expectedly across diagnostic groups, with UC performing the best followed by MCI and AD across all measures. Tables A3–A15 provide detailed normative data on performance by group, stratified by age, education, sex, and modality (including only administrations rated as very valid). Linear regression analyses covarying for age, sex, race, education, and depression revealed significant differences across diagnostic groups on each T-Cog measure in expected directions, with medium to large effect sizes ( $9.1 \leq F(2,164) \leq 186.2$ ,  $0.11 \leq \eta^2 \leq 0.70$ ,  $p$ 's < 0.001).<sup>42</sup> Figure 4



**FIGURE 3** Contributors to poor validity of T-Cog administrations.

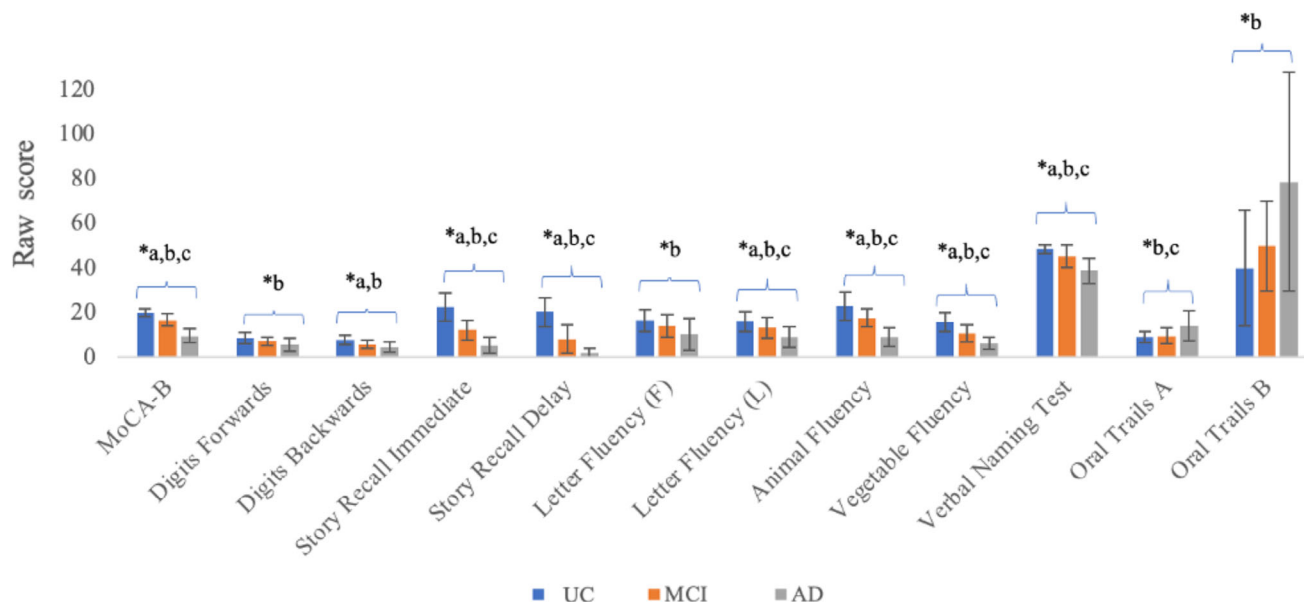
displays average raw scores on each T-Cog measure by diagnostic group across all modalities (including only those rated as very valid) to facilitate interpretability. Results were consistent when including and excluding those rated as questionably valid or invalid. When stratified by modality, all between-group differences remained consistent and statistically significant; minor differences in effect size across modality were observed on certain tests (e.g., Letter Fluency F, Oral Trails A), and overall effect sizes were relatively lower on Digits Forwards and Oral Trails B, although effect sizes for all tests remained medium to large (see Table A16).

### 3.3 | Concurrent validity and reliability in asymptomatic participants

A total of 153 individuals who completed the T-Cog at T3 also completed the in-person UDS at T2 and were identified as having UC at both timepoints. Table A17 includes features of this UC group across each timepoint, including only those rated as very valid (although results were consistent when all were included). Related-samples Wilcoxon signed-rank tests revealed a significant increase in depressive symptoms on the GDS between T2 and T3.

Spearman's correlations between in-person (T2) versus remote (T3) UDS scores revealed significant and moderate-level associations for tests with equivalent score ranges ( $0.51 \leq r_s \leq 0.69$ ,  $p$ -values  $< 0.01$ ). Correlations between tests that differed in format and score range were weak, yet remained significant (e.g.,  $r_s = .26$ ,  $p < 0.01$  for Written versus Oral Trails A; see Figures SB1–SB2 for scatter plots and correlation coefficients for all tests). ICCs for T2 (in-person) to T3 (remote) fell in the same range (moderate reliability,  $p$ 's  $< 0.001$ ) as ICCs for T1 (in-person) to T2 (in-person) for all tests except Story Recall Delay and those with different score ranges (i.e., Written vs Oral Trails, MINT vs verbal naming test), which unsurprisingly showed poor reliability from  $T2 > T3$ ; see Table A18).

As depicted in Figure SC1, average scores on UDS tests with equivalent score ranges were comparable at T2 (in-person) and T3 (remote), without significant differences according to repeated-measures ANOVA. As hypothesized, LMMs did not reveal a significant effect of timepoint (T2 vs T3) as a primary predictor of any UDS test score with equivalent score ranges (all  $p$ -values  $> 0.05$  for timepoint fixed effect), suggesting that the remote format of test administration at T3 did not meaningfully impact performance among UCs. LMMs also did not reveal a significant effect of modality (telephone vs videoconference) as a fixed effect, and the interaction between modality and



\*Linear regression models covarying for age, sex, race, education and depression is significant ( $p < .01$ )

Note: Error bars reflect standard deviations.

Note: Bars depict average raw test scores from T-Cog administrations rated as “very valid” across all modalities; between-group differences are consistent when including questionably valid/invalid administrations and within each modality (telephone vs. videoconference).

a UC vs. MCI; Bonferroni-adjusted pairwise comparison ( $p < .05$ )

b UC vs. AD; Bonferroni-adjusted pairwise comparison ( $p < .05$ )

c MCI vs. AD; Bonferroni-adjusted pairwise comparison ( $p < .05$ )

MoCA-B: Montreal Cognitive Assessment-Blind; UC: Unimpaired cognition; MCI: Mild Cognitive Impairment; AD: dementia due to Alzheimer’s disease

**FIGURE 4** Performance on T-Cog measures by diagnostic group—all modalities.

timepoint was significant for Digits Forwards only (estimate =  $-0.74$ , Standard Error (SE) =  $.37$ ,  $p = 0.048$ ); however, the 95% confidence interval (CI;  $-1.65$  to  $0.12$ ) spans zero, suggesting a nonsignificant difference. Given the overall minimal impact of modality across all preceding analyses, subsequent analyses were not stratified by modality to maximize statistical power and interpretability.

### 3.4 | Concurrent validity in all groups

A subset of individuals ( $N = 125$ ) completed UDS testing across three consecutive timepoints (T1, T2, T3). In the UC group ( $n = 91$ ), performance was relatively consistent over time (Figure 5A). Repeated-measures ANOVA revealed a significant difference between the change score of the letter fluency (L) measure only between T1 → T2 and T2 → T3 [ $F(1,84) = 5.54$ ,  $p = 0.021$ ], although this difference (0.07 points, or 0.016 SD) was not clinically significant.

In the MCI group ( $n = 15$ ), performance was variable over time, with some tests declining gradually and others improving from T1 → T2 but declining at T3 (Figure 5B). Repeated-measures ANOVA revealed a sig-

nificant difference between the T1 → T2 and T2 → T3 change score for Animal Fluency only [ $F(1,9) = 6.90$ ,  $p = 0.027$ ]. Inspection of average Animal Fluency performance at each timepoint revealed a slightly greater decline from T1 → T2 than T2 → T3, although this difference is not clinically significant (0.33 points, or 0.071 SD).

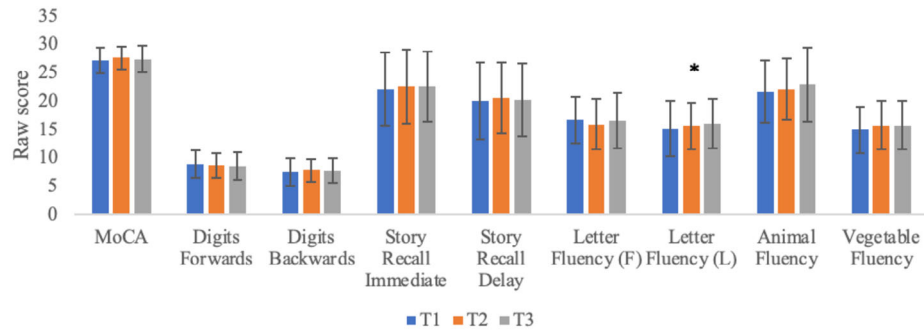
In the AD group ( $n = 19$ ), performance on most tests declined gradually over time with some exceptions (Figure 5C). Repeated-measures ANOVA revealed a significant difference between the T1 → T2 and T2 → T3 change scores for MoCA [ $F(1,12) = 6.40$ ,  $p = 0.026$ ] and Digits Forwards [ $F(1,12) = 9.15$ ,  $p = 0.011$ ]. There was a slightly greater rate of decline on both the MoCA and Digits Forwards from T2 → T3 versus T1 → T2 (by 0.54 and 0.52 points, or 0.13 SD and 0.21 SD, respectively). Again, these differences were not clinically significant.

## 4 | DISCUSSION

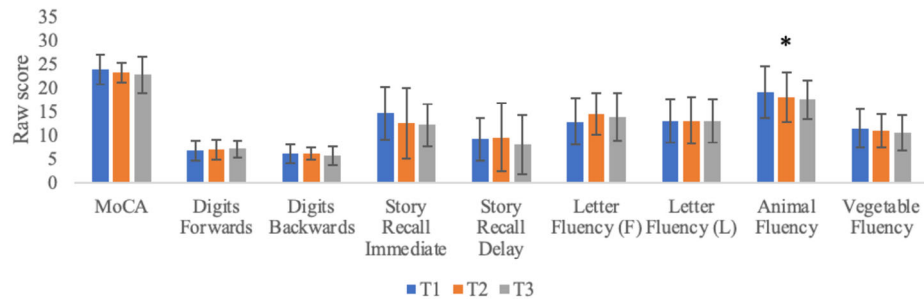
The COVID-19 pandemic necessitated a transition from in-person to remote collection of longitudinal data in clinical research cohorts and clinical practice. At the Penn ADRC, more than 200 participants



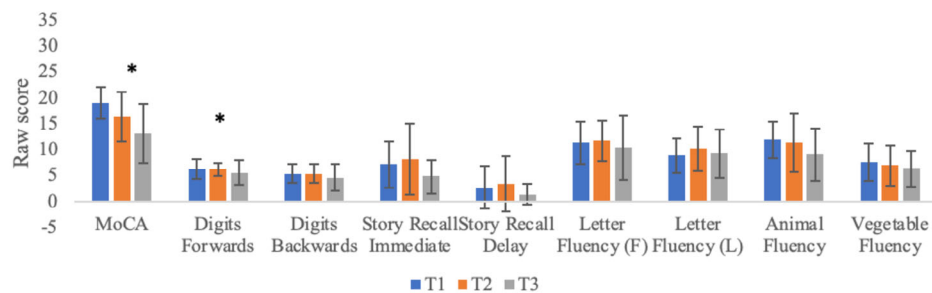
## (A) UC Group



## (B) MCI Group



## (C) AD Group



\*Repeated-measures ANOVA of T1>T2 vs. T2>T3 change scores covarying for age at T2, sex, race, education and average depression is significant ( $p<.05$ ).

Note: T1 and T2 represent in-person UDS timepoints; T3 represents remote T-Cog.

Note: Bars reflect average raw test score at each timepoint; error bars reflect standard deviations.

MoCA-B: Montreal Cognitive Assessment-Blind; UC: Unimpaired cognition; MCI: Mild Cognitive Impairment; AD: dementia due to Alzheimer's disease

**FIGURE 5** Cognitive performance on the UDS across all timepoints by diagnostic group. UDS, Uniform Data Set.

completed the T-Cog, a telemedicine version of the UDS neuropsychological battery, as of September 2021. Many participants had completed the in-person UDS at earlier timepoints. This presented an ideal opportunity to examine the feasibility and validity of the remote T-Cog battery among a well-characterized group of older adults across the cognitive continuum—albeit in an observational, nonexperimental design. Our study includes a diagnostically diverse sample of participants with UC, MCI, and AD and contributes to a growing literature on remote neuropsychological assessment. In addition, it provides normative data for the T-Cog in demographically comparable samples.

Overall, the findings support the feasibility of administering the T-Cog in a diagnostically diverse sample. Of those approached to complete the T-Cog for their annual visit, only 2% declined due to unwillingness to engage in remote testing. This is lower than the 25% who declined remote telephone testing at other ADRCs during the early stages of the COVID-19 pandemic<sup>1</sup>; however, it is important to note that comparative samples had greater demographic diversity than the present sample, suggesting that engagement rates may differ depending on sample characteristics. Of interest, participants in our cohort who elected to complete the T-Cog via telephone versus videoconference modality had lower education and were more likely

to be non-White, Hispanic/Latinx, and female. Thus, consistent with the suggestions of other researchers, multiple modalities of administration should be offered to enhance flexibility and access.<sup>1,2</sup> The decision to offer multiple modalities is further supported by our findings that feasibility and validity outcomes did not meaningfully differ across telephone versus videoconference modality.

Across all modalities, administration time for the T-Cog was 54 min on average, with cognitively unimpaired older adults completing the T-Cog in less time than those with MCI or AD. Completion rates also differed by diagnostic group, with a lower proportion of participants with AD completing all 12 measures (26%). Most administrations were rated as valid regardless of modality—consistent with Sachs et al. (2024)<sup>4</sup>—and validity ratings were not impacted by sociodemographic factors such as age, education, race, or ethnicity. On the other hand, lower cognitive and functional status was associated with decreased likelihood of valid administration. Common contributors to poor validity included hearing impairment, connectivity difficulties, and distractions, consistent with those reported in other settings.<sup>2</sup> Hearing difficulties were particularly common among participants with MCI and AD. These findings suggest that those engaging in teleneuropsychology can proactively address potential modifiable barriers such as ensuring that participants are wearing hearing aids, speaking loudly and clearly, and using a landline to avoid connectivity difficulties. Nonetheless, longer completion times, lower completion rates, and lower validity ratings among the AD group suggest that the remote T-Cog is relatively less feasible among those with dementia-level cognitive impairment.

Preliminary validity of the T-Cog was supported through significant predicted differences and medium to large effect sizes between the UC, MCI, and AD groups on all T-Cog measures, regardless of modality, suggesting that the T-Cog can accurately measure differences in cognitive ability between diagnostic groups. Among UC, performance on the T-Cog was comparable to prior in-person UDS performance, and most tests with equivalent score ranges demonstrated relatively lower yet comparable, moderate reliability across all three timepoints, providing additional evidence for concurrent validity and adequate reliability across in-person versus remote formats. These results are consistent with studies noting moderate test-retest reliability of T-Cog measures at shorter retest intervals, in the context of moderate reliability for face-to-face administrations.<sup>3</sup> Among all diagnostic groups, rates of change across the three timepoints followed expected trends, with stability in the UC group, variability in the MCI group, and gradual decline in the AD group.

Of note, select tests were more closely related to their in-person correlates (e.g., Digit Span, verbal fluency) than those with differences in score range and underlying cognitive mechanism (e.g., Oral vs Written Trail Making test). This is consistent with prior literature and should be considered when using tests that do not have a closely matched telemedicine version.<sup>4,43</sup> Another weakness of the T-Cog is the lack of visuospatial tasks. Ultimately, clinicians must decide whether a given remote tool is appropriate for a particular person and weigh the pros, cons, and ethical considerations carefully—particularly if results will be used for clinical purposes.<sup>21</sup> This includes balancing the benefits

of increased accessibility conferred by at-home telephone and video formats with the potential for incomplete cognitive characterization, which has the potential to exacerbate or mitigate pre-existing higher rates of missed diagnosis in individuals from socioeconomically disadvantaged groups, depending on how telemedicine-based evaluations are applied.<sup>44</sup>

Several studies have reported limited access to requisite technology and poor confidence using technology as barriers to engaging in teleneuropsychology.<sup>2,21,45,46</sup> It is encouraging that a dual phone-and/or video-based protocol such as the T-Cog can circumvent these barriers, and that participants in other cohorts have reported high levels of satisfaction with and preference for remote over in-person assessment.<sup>4</sup> Nonetheless, others have reported missing the element of human contact with research staff during in-person visits,<sup>4</sup> with lower satisfaction rates for telephone assessment among older adults facing loneliness, depression, and isolation.<sup>47</sup> Another drawback to the remote T-Cog is that it is currently unavailable in languages other than English and Spanish, which limits accessibility. Development and validation studies in community-based, demographically diverse, non-English-speaking populations are needed to ensure existing health care disparities are not widened. A major limitation of the current study is that although there was more representation in the UC group, the symptomatic groups were largely White, non-Hispanic, and highly educated, which limits generalizability. Underrepresentation of minoritized groups in AD/ADRD trials and research is an urgent issue<sup>48,49</sup> and the Penn ADRC has several ongoing efforts for engaging and partnering with local communities which have led to increased cohort diversity.

Taken together, these results suggest that using the T-Cog for remote neuropsychological assessment is a feasible and comparable—albeit nonequivalent—alternative to in-person testing and provides a valid characterization of cognitive functioning across the diagnostic continuum. The T-Cog may be compared longitudinally to in-person testing and allows for increased access to clinical and research services, reduces missing data points in longitudinal cohorts, and encourages retention of participants. Nonetheless, the results suggest that longitudinal comparisons should be interpreted with caution for select subtests, such as the Oral Trails A/B. A major limitation of the current study is its observational design, which precluded experimental control over order of in-person versus T-Cog administrations and time intervals between assessments. Current ongoing studies will extend these results in a well-controlled counter-balanced design.<sup>50</sup> Other future directions include ongoing validation to track clinical progression and compare the discriminative accuracy of the T-Cog to the in-person UDS, expansion of normative data for the Spanish T-Cog, combining normative data across ADRC sites, and expanding validation efforts in more socioculturally diverse populations.

## ACKNOWLEDGMENTS

This study is supported by the National Institutes of Health (NIH) and National Institute on Aging (NIA), including the following grants: F31AG069444 and T32AG066598 (Katherine Hackett); P30AG072979; K23AG065499 (Dawn Mechanic-Hamilton).

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest. Author disclosures are available in the [supporting information](#).

## CONSENT STATEMENT

All individuals provided written informed consent approved by the institutional review board at the University of Pennsylvania prior to undergoing any study procedures or assessments as part of their participation in the Penn ADRC Clinical Core cohort.

## ORCID

Katherine Hackett  <https://orcid.org/0000-0002-3595-1418>

Dawn Mechanic-Hamilton  <https://orcid.org/0000-0002-8728-2420>

## REFERENCES

- Loizos M, Neugroschl J, Zhu CW, et al. Adapting Alzheimer disease and related dementias clinical research evaluations in the age of COVID-19. *Alzheimer Dis Assoc Disord*. 2021;35(2):172. doi:10.1097/WAD.0000000000000455
- Fox-Fuller JT, Rizer S, Andersen SL, Sunderaraman P. Survey findings about the experiences, challenges, and practical advice/solutions regarding teleneuropsychological assessment in adults. *Arch Clin Neuropsychol*. 2022;37(2):274-291. doi:10.1093/arclin/acab076
- Howard RS, Goldberg TE, Luo J, Munoz C, Schneider LS. Reliability of the NACC telephone-administered neuropsychological battery (T-cog) and Montreal cognitive assessment for participants in the USC ADRC. *Alzheimers Dement*. 2023;15(1):e12406. doi:10.1002/dad2.12406
- Sachs BC, Latham LA, Bateman JR, et al. Feasibility of remote administration of the uniform data set-version 3 for assessment of older adults with mild cognitive impairment and Alzheimer's disease. *Arch Clin Neuropsychol*. 2024;39(5):635-643. doi:10.1093/arclin/aca001
- Kane RL, Parsons TD. *The Role of Technology in Clinical Neuropsychology*. Oxford University Press; 2017.
- Bilder RM, Postal KS, Barisa M, et al. Inter organizational practice committee recommendations/guidance for teleneuropsychology (TeleNP) in response to the COVID-19 pandemic. *Clin Neuropsychol*. 2020;34(7-8):1314-1334. doi:10.1080/13854046.2020.1767214
- Chapman JE, Gardner B, Ponsford J, Cadilhac DA, Stolwyk RJ. Comparing performance across in-person and videoconference-based administrations of common neuropsychological measures in community-based survivors of stroke. *J Int Neuropsychol Soc*. 2021;27(7):697-710. doi:10.1017/S1355617720001174
- Cullum CM, Grosch MC. Special considerations in conducting neuropsychology assessment over videoteleconferencing. In: Myers K, Turvey CL, eds. *Telemental Health: Clinical, Technical, and Administrative Foundations for Evidence-Based Practice*. Elsevier; 2013:275-293.
- Cullum CM, Hynan LS, Grosch M, Parikh M, Weiner MF. Teleneuropsychology: evidence for video teleconference-based neuropsychological assessment. *J Int Neuropsychol Soc*. 2014;20(10):1028-1033. doi:10.1017/S1355617714000873
- Hildebrand R, Chow H, Williams C, Nelson M, Wass P. Feasibility of neuropsychological testing of older adults via videoconference: implications for assessing the capacity for independent living. *J Telemed Telecare*. 2004;10(3):130-134.
- Wadsworth HE, Dhima K, Womack KB, et al. Validity of teleneuropsychological assessment in older patients with cognitive disorders. *Arch Clin Neuropsychol*. 2018;33(8):1040-1045. doi:10.1093/arclin/acx140
- Brearily TW, Shura RD, Martindale SL, et al. Neuropsychological test administration by videoconference: a systematic review and meta-analysis. *Neuropsychol Rev*. 2017;27(2):174-186. doi:10.1007/s11065-017-9349-1
- Marra DE, Hamlet KM, Bauer RM, Bowers D. Validity of teleneuropsychology for older adults in response to COVID-19: A systematic and critical review. *Clin Neuropsychol*. 2020;34(7-8):1411-1452. doi:10.1080/13854046.2020.1769192
- Postal KS, Bilder RM, Lanca M, et al. Inter organizational practice committee guidance/recommendation for models of care during the novel coronavirus pandemic. *Arch Clin Neuropsychol*. 2021;36(1):17-28. doi:10.1093/arclin/acia073
- Chapman JE, Ponsford J, Bagot KL, Cadilhac DA, Gardner B, Stolwyk RJ. The use of videoconferencing in clinical neuropsychology practice: A mixed methods evaluation of neuropsychologists' experiences and views. *Aust Psychol*. 2020;55(6):618-633. doi:10.1111/ap.12471
- Marra DE, Hoelzle JB, Davis JJ, Schwartz ES. Initial changes in neuropsychologists clinical practice during the COVID-19 pandemic: a survey study. *Clin Neuropsychol*. 2020;34(7-8):1251-1266. doi:10.1080/13854046.2020.1800098
- Fox-Fuller JT, Ngo J, Pluim CF, et al. Initial investigation of test-retest reliability of home-to-home teleneuropsychological assessment in healthy, English-speaking adults. *Clin Neuropsychol*. 2021;36(8):2153-2167. doi:10.1080/13854046.2021.1954244
- Parks AC, Davis J, Spreser CD, Stroescu I, Ecklund-Johnson E. Validity of in-home teleneuropsychological testing in the wake of COVID-19. *Arch Clin Neuropsychol*. 2021;36(6):887-896. doi:10.1093/arclin/acab002
- Parsons MW, Gardner MM, Sherman JC, et al. Feasibility and acceptance of direct-to-home tele-neuropsychology services during the COVID-19 pandemic. *J Int Neuropsychol Soc*. 2022;28(2):210-215. doi:10.1017/S1355617721000436
- Sperling SA, Acheson SK, Fox-Fuller J, et al. Tele-neuropsychology: from science to policy to practice. *Arch Clin Neuropsychol*. 2024;39(2):227-248. doi:10.1093/arclin/acad066
- Scott TM, Marton KM, Madore MR. A detailed analysis of ethical considerations for three specific models of teleneuropsychology during and beyond the COVID-19 pandemic. *Clin Neuropsychol*. 2022;36(1):24-44. doi:10.1080/13854046.2021.1889678
- Weintraub S, Besser L, Dodge HH, et al. Version 3 of the Alzheimer Disease Centers' Neuropsychological test battery in the Uniform Data Set (UDS). *Alzheimer Dis Assoc Disord*. 2018;32(1):10-17. doi:10.1097/WAD.0000000000000223
- Abdolahi A, Bull MT, Darwin KC, et al. A feasibility study of conducting the Montreal Cognitive Assessment remotely in individuals with movement disorders. *Health Informatics J*. 2016;22(2):304-311. doi:10.1177/1460458214556373
- Nasreddine ZS, Phillips NA, Bedirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53(4):695-699. doi:10.1111/j.1532-5415.2005.53221.x
- Craft S, Newcomer J, Kanne S, et al. Memory improvement following induced hyperinsulinemia in Alzheimer's disease. *Neurobiol Aging*. 1996;17(1):123-130. doi:10.1016/0197-4580(95)02002-0
- Yochim BP, Beaudreau SA, Fairchild JK, et al. Verbal naming test for use with older adults: development and initial validation. *J Int Neuropsychol Soc*. 2015;21(3):239-248. doi:10.1017/S1355617715000120
- Ricker JH, Axelrod BN. Analysis of an oral paradigm for the trail making test. *Assessment*. 1994;1(1):47-51. doi:10.1177/1073191194001001007
- Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clin Gerontol*. 1986;5:165-173.
- Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The neuropsychiatric inventory. *Neurology*. 1994;44(12):2308. doi:10.1212/WNL.44.12.2308

30. Clark CM, Ewbank DC. Performance of the dementia severity rating scale: a caregiver questionnaire for rating severity in Alzheimer disease. *Alzheimer Dis Assoc Disord.* 1996;10(1):31.
31. Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *J Gerontol.* 1982;37(3):323-329. doi:10.1093/geronj/37.3.323
32. Morris JC. The clinical dementia rating (CDR) current version and scoring rules. *Neurology.* 1993;43(11):2412-2412.
33. Byrd DA, Rivera-Mindt MG. Neuropsychology's race problem does not begin or end with demographically adjusted norms. *Nat Rev Neurol.* 2022;18(3):125-126. doi:10.1038/s41582-021-00607-4
34. Chan KCG, Barnes LL, Saykin AJ, et al. Racial-ethnic differences in baseline and longitudinal change in neuropsychological test scores in the NACC Uniform Data Set 3.0. *Alzheimers Dement.* 2021;17(S6):e054653. doi:10.1002/alz.054653
35. Fyffe DC, Mukherjee S, Barnes LL, Manly JJ, Bennett DA, Crane PK. Explaining differences in episodic memory performance among older African Americans and Whites: the roles of factors related to cognitive reserve and test bias. *J Int Neuropsychol Soc.* 2011;17(4):625-638. doi:10.1017/S1355617711000476
36. Manly JJ, Byrd DA, Touradjji P, Stern Y. Acculturation, reading level, and neuropsychological test performance among African American elders. *Appl Neuropsychol.* 2004;11(1):37-46. doi:10.1207/s15324826an1101\_5
37. McDermott LM, Ebmeier KP. A meta-analysis of depression severity and cognitive function. *J Affect Disord.* 2009;119(1):1-8. doi:10.1016/j.jad.2009.04.022
38. Sachs BC, Steenland K, Zhao L, et al. Expanded demographic norms for version 3 of the Alzheimer Disease Centers' neuropsychological test battery in the Uniform Data Set. *Alzheimer Dis Assoc Disord.* 2020;34(3):191-197. doi:10.1097/WAD.0000000000000388
39. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15(2):155-163. doi:10.1016/j.jcm.2016.02.012
40. MoCA Test Inc. MoCA Cognition FAQ. Remote MoCA Testing. MoCA Cognition; 2023. <https://mocacognition.com/faq/>
41. Byrom B, Breendon P, Tulkki-Wilke R, Platko J. Meaningful change: defining the interpretability of changes in endpoints derived from interactive and mHealth technologies in healthcare and clinical research. *J Rehabil Assist Technol Eng.* 2020;7:2055668319892778. doi:10.1177/2055668319892778
42. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* 2nd ed. Erlbaum; 1988.
43. Fox-Fuller JT, Vickers KL, Saurman JL, Wechsler R, Goldstein FC. Comparison of performance on the oral versus written trail making test in patients with movement disorders. *Appl Neuropsychol Adult.* 2023:1-7. doi:10.1080/23279095.2023.2288230
44. Mattke S, Jun H, Chen E, Liu Y, Becker A, Wallick C. Expected and diagnosed rates of mild cognitive impairment and dementia in the U.S. Medicare population: observational analysis. *Alzheimers Res Ther.* 2023;15(1):128. doi:10.1186/s13195-023-01272-z
45. Rochette AD, Rahman-Filipiak A, Spencer RJ, Marshall D, Stelmokas JE. Teleneuropsychology practice survey during COVID-19 within the United States. *Appl Neuropsychol Adult.* 2022;29(6):1312-1322. doi:10.1080/23279095.2021.1872576
46. Loizos M, Baim-Lance A, Ornstein KA, et al. If you give them away, it still may not work: challenges to video telehealth device use among the urban homebound. *J Appl Gerontol.* 2023;42(9):1896-1902. doi:10.1177/07334648231170144
47. Sewell M, Neugroschl J, Zhu CW, et al. Lessons from the coronavirus disease experience: research participant and staff satisfaction with remote cognitive evaluations. *Alzheimer Dis Assoc Disord.* 2024;38(1):65. doi:10.1097/WAD.0000000000000605
48. Manly JJ, Glymour MM. What the aducanumab approval reveals about Alzheimer disease research. *JAMA Neurol.* 2021;78(11):1305-1306. doi:10.1001/jamaneurol.2021.3404
49. Raman R, Aisen P, Carillo MC, et al. Tackling a major deficiency of diversity in Alzheimer's disease therapeutic trials: an CTAD Task Force Report. *J Prev Alzheimers Dis.* 2022;9(3):388-392. doi:10.14283/jpad.2022.50
50. Sachs BC, Latham L, Craft S, et al. Technology use amongst older adults with and without cognitive impairment: results from the 'Validation of Video Administration of a Modified UDSv3 Cognitive Battery (VCog)' Study. *Alzheimer's Association International Conference.* Alzheimer's Association; 2024.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Hackett K, Shi Y, Schankel L, et al. Feasibility, validity, and normative data for the remote Uniform Data Set neuropsychological battery at the University of Pennsylvania Alzheimer's Disease Research Center. *Alzheimer's Dement.* 2024;16:e70043. <https://doi.org/10.1002/dad2.70043>