

Research Article

Videoconference-integrated, computer-assisted cognitive testing improves the remote assessment of processing speed and attention

Jodie E. Chapman¹ , Christoph Helmstaedter² , David F. Abbott^{1,3,4} , Heath R. Pardoe^{1,3} , David N. Vaughan^{1,3,5} , Graeme D. Jackson^{1,3,5} and Chris Tailby^{1,3,6} for the Australian Epilepsy Project Investigators^{*}

¹Florey Institute of Neuroscience and Mental Health, Heidelberg, VIC, Australia, ²Department of Epileptology, University of Bonn (UKB), Bonn, Germany, ³Florey Department of Neuroscience and Mental Health, The University of Melbourne, Australia, ⁴Department of Medicine, Austin Health, The University of Melbourne, Melbourne, Australia, ⁵Department of Neurology, Austin Hospital, Heidelberg, Australia and ⁶Department of Clinical Neuropsychology, Austin Hospital, Heidelberg, Australia

Abstract

Objective: Remote videoconference neuropsychological assessments offer opportunities that remain under-exploited. We aimed to evaluate teleneuropsychology (TeleNP)-suitable oral and digital versions of the Symbol Digit Modalities Task (SDMT) and Trail Making Test (TMT) – widely used measures of speed and attention – by comparing them to their written counterparts. **Methods:** Three-hundred and twenty-one Australian Epilepsy Project (AEP) adult participants with seizure disorders completed the written SDMT and TMT in-person. One-hundred and forty-four of these participants also completed the oral SDMT and TMT during a remote videoconference-based assessment while 177 completed a novel, examiner-administered digital SDMT analogous measure named Symbol Decoding and a novel digital TMT remotely via custom videoconference-based software. **Results:** Oral SDMT and digital Symbol Decoding strongly correlated with in-person written SDMT (r(133) = .77, p < .001 and r(126) = .76, p < .001, respectively). Oral TMT-B was only moderately associated (r(126) = .52, p < .001) with written TMT-B and, less strongly related to measures of sustained attention and spatial working memory than its written counterpart. Digital TMT better reproduced the written test's properties with improved association with written TMT-B (r(154) = .71, p < .001). **Conclusions:** Oral SDMT and digital Symbol Decoding are strongly correlated with in-person written SDMT. The digital TMT better captures the cognitive demands and performance characteristics of the in-person written form than does oral TMT. Videoconference-integrated digital tasks offer increased standardization and automation in administration and scoring and the potential for rich metadata, making them an attractive area for further development.

Keywords: attention; neuropsychological tests; information processing speed; symbol digits modalities task; telehealth; trail making test (Received 29 October 2024; final revision 7 May 2025; accepted 8 June 2025)

Statement of Research Significance

Research Question(s) or Topic(s): Teleneuropsychology has increased in popularity in recent years. Processing speed and attention deficits are common neuropsychological findings across conditions, however, traditional paper-and-pen measures assessing these domains are difficult to administer via telehealth. We evaluated novel, videoconference-suitable adaptations of the Symbol Digit Modalities Task and Trail Making Test – popular measures of speed and attention – by comparing them to their written counterparts. Main Findings: The oral Symbol Digit

Modalities Task and our novel digital Symbol Decoding task both strongly correlated with the written Symbol Digital Modalities Task. Oral Trail Making Test – B was only moderately associated with the written Trail Making Test – B, however, our novel digital Trail Making Test better reproduced the written task's properties. **Study Contributions:** We demonstrate that these novel, videoconference-integrated digital tasks are comparable to their traditional written counterparts and offer increased standardization, automation, and rich metadata, making them a promising area for further development.

Corresponding author: Chris Tailby; Email: chris.tailby@florey.edu.au

*A full list of Australian Epilepsy Project Investigators with a Contribution Roles Taxonomy (CRediT) author statement for this manuscript is available in Supplementary Materials. Cite this article: Chapman J.E., Helmstaedter C., Abbott D.F., Pardoe H.R., Vaughan D.N., Jackson G.D., & Tailby C. for the Australian Epilepsy Project Investigators. Videoconference-integrated, computer-assisted cognitive testing improves the remote assessment of processing speed and attention. *Journal of the International Neuropsychological Society*, 1–10, https://doi.org/10.1017/S1355617725101124

© The Author(s), 2025. Published by Cambridge University Press on behalf of International Neuropsychological Society. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Introduction

The use of teleneuropsychology (TeleNP) has steadily grown in recent years, particularly since the coronavirus pandemic (COVID-19; e.g., Kitaigorodsky et al., 2021; Tailby et al., 2020; Zane, Thaler, Reilly, Mahoney, & Scarisbrick, 2021). TeleNP can address service access barriers (e.g., a scarcity of clinicians in rural/remote locations, mobility and/or driving restrictions) and also offers additional clinical benefits (e.g., the capacity to see clients in their own home; Chapman et al., 2020). TeleNP can also enhance research by enabling centrally collected nationwide data, which can improve data quality and diversity. To date, however, the dominant approach to TeleNP assessment has been to administer traditional neuropsychological measures on a videoconference call, without capitalizing on the TeleNP modality itself (e.g., using the technology to facilitate task administration and scoring).

Processing speed and attention deficits are among the most common neuropsychological findings across neurological and psychological conditions (e.g., mood disorders: Marvel & Paradiso, 2004; stroke: Nys et al., 2007; multiple sclerosis [MS]: Oreja-Guevara et al., 2019; traumatic brain injury [TBI]: Ponsford et al., 2013, 2014; epilepsy: van Rijckevorsel, 2006). Theses domains are, however, particularly difficult to assess via TeleNP because evidence-based measures assessing these domains, such as the Symbol Digit Modalities Task (SDMT) and Trail Making Test (TMT), are written tasks in which participants record their responses on a paper form(s) so they are not easily administered over a videoconference call. The written SDMT is a timed symbol transcription task in which examinees decode symbols via a provided key; it is primarily a measure of psychomotor and cognitive processing speed and attention (Smith, 1991). The written TMT is a timed two-part join-the-dots task in which examinees draw a line connecting (1 - 25) in ascending order (TMT-A) and then numbers (1 -13) and letters (A – L) in alternating ascending order (TMT-B). TMT-A indexes psychomotor and cognitive processing speed and visual search/visual working memory, while TMT-B adds additional attention and executive elements to the task, such as attentional switching (Strauss et al., 2006).

Both the SDMT and TMT have oral versions that can, and have, been administered via TeleNP (e.g., Eilam-Stock et al., 2021). The oral SDMT requires examinees to provide verbal rather than written responses (Smith, 1991). Thus, the demands of the written and oral SDMT are comparable, as the tasks differ only in response modality (i.e., written versus spoken). Existing non-TeleNP administered research confirms this in both healthy (e.g., r > .78; Smith, 1991) and clinical samples (e.g., TBI: r = .88; Ponsford & Kinsella, 1992; MS: r = .89; Sandroff et al., 2013).

The oral TMT, however, somewhat departs from its written form, requiring only that examinees orally recite the sequences they would otherwise search for and mark out in the written TMT (Ricker & Axelrod, 1994). Consequently, the demands of the written and oral forms of the TMT diverge (Bastug et al., 2013; Jaywant et al., 2018; Mrazik et al., 2010). In its written form, the path established by completing the task serves as a visual reminder of the underlying sequence and can be used to minimize the demands on divided attention (Bastug et al., 2013), a support unavailable in the oral form. The written form also requires visual search, engages visuospatial working memory, and takes longer than the oral form, thereby placing greater demands on sustained attention and being more sensitive to slowed processing speed. Observed correlations between the written and oral TMT-B have

varied from weak to moderate (Bastug et al., 2013: r = .69; Grigsby & Kaye, 1995: r = .38; Mrazik et al., 2010: r = .62; Ricker & Axelrod, 1994: r = .72).

These oral SDMT and TMT do not capitalize on the fact that the intrinsic reliance upon a computer-assisted medium for TeleNP assessment also presents opportunities to improve the ease, efficiency, and utility of cognitive testing (Tailby et al., 2024). Accordingly, several digital adaptations of the SDMT and TMT have been evaluated. Evidence suggests that digital versions of the SDMT are reliable, valid, and sensitive indicators of impairment in various clinical populations (e.g., Akbar et al., 2011; Bigi et al., 2017; Forn et al., 2009; Hardy et al., 2021), although few direct comparisons have been made between digital SDMTs and oral and/or written versions of this task. Similarly, evidence indicates that digital TMTs measure similar constructs to written TMT (e.g., Baykara et al., 2022; Dahmen et al., 2017; Fellows et al., 2017; Lunardini et al., 2019; Park & Schott, 2022) and capture essential elements of the written TMT better than the oral version (e.g., visual search, a visual path that must be tracked, longer completion times). To our knowledge, however, no videoconference-enabled digital SDMT and/or TMT have been developed. Videoconference-enabled digital tasks offer the benefit of clinician-administered digital tasks but also extend the benefits to include those of remote TeleNP broadly (e.g., improved geographical reach, reduced travel).

The overarching aim of this study was to compare TeleNPsuitable versions of the SDMT and TMT (i.e., remotely administered oral and digital versions) with the sensitive and frequently used written versions of these tasks. That is, our primary research question is one of clinical and practical relevance – given the practical constraints imposed by a remote TeleNP context, which of the TeleNP-appropriate test forms (oral or digital) yields results comparable to the in-person administered written formants? Accordingly, the two specific aims of this study were to (a) compare the remotely administered *oral* SDMT and TMT to the in-person administered written SDMT and TMT, respectively, and (b) to compare remotely administered novel digital versions of the SDMT (named Symbol Decoding) and TMT to the in-person administered written SDMT and TMT, respectively. If there was a weaker correlation between task versions than anticipated, an additional aim was to evaluate the underlying basis of the weaker relationship by evaluating how each task version correlated with measures of other cognitive constructs. It is important to address these aims by evaluating a cohort that includes individuals with cognitive impairment, so as to capture task comparability across a broad range of ability levels. To this end, we evaluated these aims in a cohort of individuals with seizure disorders, in whom processing speed and attention deficits are common (van Rijckevorsel, 2006). Furthermore, the written SDMT and TMT have been shown to be sensitive to medication effects in epilepsy (Lutz & Helmstaedter, 2005). As such, those with seizure disorders represent a useful clinical sample in which to evaluate these aims.

Given that the written SDMT, oral SDMT, and digital Symbol Decoding differ only in the response modality required, we hypothesized that both the remotely administered oral SDMT and remotely administered digital Symbol Decoding would have strong positive correlations with the in-person administered written SDMT. In contrast, as the written and oral TMT have different task demands, we hypothesized there would be only a modest relationship between the in-person written and remotely administered oral TMT. However, as the remotely administered digital TMT more closely aligns with its written counterpart, we

hypothesized that the written and digital TMT would be strongly correlated, and, indeed, more strongly correlated than the oral and written TMT.

Method

Design

A within-subjects non-randomized design was employed. We evaluated two separate samples collected in the Australian Epilepsy Project (AEP; see https://epilepsyproject.org.au/) to address each aim in turn. To address Aim One, AEP participants completed the oral SDMT and TMT during their AEP TeleNP session (see Tailby et al., 2020 for a description of the TeleNP procedures used in the pilot AEP). A protocol change following the pilot AEP resulted in the oral SDMT and TMT being replaced with novel digital versions of these tasks, administered via custom TeleNP software. To address Aim Two, later recruited AEP participants completed digital Symbol Decoding and digital TMT in their AEP TeleNP session via the custom TeleNP software. For both samples, the written SDMT and TMT were completed during an in-person site visit at which their AEP Magnetic Resonance Imaging (MRI) scans were acquired.

Participants

AEP participants are adults aged between 18 and 67 years with either (a) a first unprovoked seizure, (b) a new diagnosis of epilepsy (within six months), or (c) pharmacoresistant focal epilepsy. AEP participants are required to have a functional level of English; exclusion criteria are a moderate or severe intellectual disability and/or contraindications for 3 tesla MRI. Study participants were referred by neurologists primarily at Austin Health and from other private or public health services in Melbourne, Australia.

Measures

Participant demographic and clinical data were obtained during initial eligibility screening, from the participants treating clinician at the time of referral, and/or from a medical history interview.

During the pilot phase of the project participants completed a battery of neuropsychological measures during an AEP TeleNP session conducted via Zoom (Zoom Video Communications, Inc., 2021). The neuropsychological measures were selected (a) to ensure assessment of epilepsy-relevant domains of cognition, and (b) based on tasks with demonstrated sensitivity in epilepsy (Lutz & Helmstaedter, 2005; Tailby et al., 2020). The subset of measures relevant to the current evaluations include the Test of Premorbid Function (ToPF; Wechsler, 2009) to derive an intelligence quotient estimate; Weschler Adult Intelligence Scale - Fourth Edition (WAIS-IV) Digit Span Backward subtest (Wechsler, 2008); Letter Fluency (Strauss et al., 2006); and the web-based CANTAB Spatial Working Memory (SWM) and Rapid Visual Processing (RVP) subtests (Cambridge Cognition, 2019). Not all participants completed the CANTAB measures due to a change in protocol after the first 100 participants completed the AEP. Neuropsychological measures were administered and scored according to standard instructions. An exception was WAIS-IV Digit Span Backward. For this task an adapted administration was used in which only the backward span items were administered, with the second item of a given length only administered if the first item was incorrect, yielding the Longest Digit Span Backwards (LDSB).

As previously described, our custom TeleNP software was introduced part way through data collection in the AEP and replaced the protocol described in Tailby et al. (2020). This allowed evaluation of *Aim Two*. Our software is custom-built video-conference software developed in-house in which task administration was guided by the examiner with standardized instructions provided on their screen. Task administration/completion was facilitated with the use of technology (e.g., responses were recorded and scored on the software) or via screen share combined with paper record forms.

SDMT and digital symbol decoding

The written and oral SDMT were administered according to the standardized instructions provided by Smith (1991). Given the oral SDMT was administered in the TeleNP setting, a digital version of the stimulus sheet was shown to the participant via the Zoom screenshare function. The Smith (1991) SDMT form was administered when conducting the written administration, and the Hinton-Bayre & Geffen (2005) Form B for the oral administration. A custom digital SDMT analogous measure named Symbol Decoding was administered via our custom TeleNP software. Like the written/oral SDMT, digital Symbol Decoding contains a key matching nine unique symbols to the numbers one to nine. The participant must decode subsequent rows of symbols in line with the provided key. Participants practiced 10 symbols to demonstrate their understanding of the task. Digital Symbol Decoding presented the key and four rows of ten symbols on each page (up to three pages in total). The examiner screen displayed the same stimulus set displayed to the participant, but with the correct responses included in the box beneath each symbol. The examiner recorded whether each response was correct or incorrect by pressing the number (shaded green) or symbol (shaded red), respectively, for each item in the response key shown on their screen, with the color of responded items bolded on the examiner screen to indicate the response was registered. Recorded responses can be corrected in real time by the examiner (e.g., if the participant subsequently self-corrects an error). The page automatically advances when the examiner has recorded the final response on a given page. The task was administered with similar instructions to the oral and written counterparts using the same 90 s time-limit. For both the oral SDMT and digital Symbol Decoding, participants were instructed to use their cursor or finger to track their progress if desired.

TMT

The written TMT was administered according to the standard administration instructions provided by Strauss et al. (2006). The oral TMT was administered according to the instructions outlined in Ricker and Axelrod (1994) and errors were corrected using the instructions provided by Mrazik et al. (2010). A custom digital TMT was administered via our custom TeleNP software. Like the written/oral TMT, participants completed a practice trial followed by a timed test trial in which they connected numbers from 1-25(TMT-A) and then numbers (1 - 13) and letters (A - L) in alternating ascending order (TMT-B). Participants responded by using their mouse to select each number in sequence. As participants responded, each correctly selected number/letter turned grey and joined to the preceding number/letter with a line to create the trail. If a participant made an error, the incorrectly selected number/letter turned red, and a message appeared at the top of the screen directing the participant back to their last correct response. The message provided increasing guidance if successive

errors were made. The examiner saw a mirrored version of the participants display and their mouse clicks/position. The examiner provided additional verbal or visual (cursor) prompts if the participant did not recognize they had made an error and/or could not be successfully redirected with the provided prompts.

Procedure

This research was completed in accordance with the Helsinki Declaration. Ethics approval for the AEP was granted by Austin Health Human Research Ethics Committee (HREC/68372/Austin-2022). Participants provided their written informed consent prior to participation.

Trained research assistants conducted the TeleNP sessions. Participants completed these sessions either at home on their own laptop/desktop ($Aim\ One, n = 136, 94.4\%; Aim\ Two, n = 161, 91\%$) or at the research site on a provided laptop/desktop computer while the research assistant conducted the session in a separate room ($Aim\ One,\ n=8,5.6\%;\ Aim\ Two,\ n=16,9\%$). The latter was the case when participants did not have an appropriate device or internet connection, or for other reasons were unable to complete the session at home. Participants were able to use any microphoneand camera-enabled laptop/desktop. They were instructed to use a USB/wireless mouse if using a laptop. Our custom TeleNP software automatically captured details of the participants' browser (name, version, window size), operating system (name, version), and monitor resolution. In each TeleNP session, the research assistant went through a checklist with participants to ensure the environment in which they were conducting the TeleNP session was appropriate. This included, for example, ensuring they were alone in a private, distraction-free environment, that computer notifications were turned off, and confirming their location, phone number, and an emergency contact in the event of a seizure. The written-form versions were administered by trained research assistants when the participants attended the research location for an MRI. For most participants, the TeleNP session occurred prior to the MRI scan. Specifically, for Aim One, 124 (86.1%) participants completed the oral tasks before the written tasks, and 20 (13.9%) completed the written tasks before the oral tasks. The median time between sessions for this sample was 4 days (IQR = 1, 9.5). For Aim Two, all participants except one completed the digital tasks prior to the written tasks. The median time between sessions for this sample was 7 days (IQR = 3, 9). For all neuropsychological measures completed in TeleNP sessions, the administering research assistant characterized the completion of the test as "complete and reliable" or impacted by other factors (e.g., distraction, refusal). These same evaluations were made retrospectively for written SDMT and TMT performances by reviewing examiners' comments made at the time of the assessment and the response forms.

Data analyses

Data analyses were conducted using RStudio Version 1.4.1106 (RStudio Team, 2021).

Data preparation

Only performances on neuropsychological tasks recorded as "complete and reliable" were included in analyses. Eleven participants were excluded from the sample addressing *Aim One* and two from the sample addressing *Aim Two* as they did not have "complete and reliable" data available for any pairwise comparison. For *Aim Two*, 177 participants completed the digital TMT and 140

completed digital Symbol Decoding, the latter being introduced later in the protocol. Across analyses, missing values were excluded using pairwise deletion. Outliers, identified via visual inspection of histograms, were deleted ($Aim\ One$: n=1 on written TMT-B; $Aim\ Two$: n=1 on digital TMT-B and written TMT-B). The total number of "complete and reliable" cases included for each comparison of interest is shown in Tables 2 to 4. Following this, histograms as well as Shapiro–Wilk tests (Shapiro & Wilk, 1965) were used to evaluate normality. Positively skewed variables (i.e., completion time on all TMT tasks) underwent logarithmic transformations.

Evaluating relationships between different versions of the SDMT and TMT

The relationships between the written and oral and written and digital SDMT and TMT task versions were evaluated using Pearson's product–moment correlation coefficients.

Evaluating the cognitive constructs measured by the written and oral TMT-B

As outlined in the Introduction, the oral TMT does not fully replicate the cognitive demands of the written TMT. We explored this for TMT-B, which we focus on here given the overlearned automatic nature of oral TMT-A and the relative clinical interest in TMT-B, by examining the degree to which the oral and written forms of TMT-B were associated with other measures of processing speed and high-level attention and executive processes, using Pearson's product—moment correlation coefficients or Spearman's rank correlation coefficients as appropriate.

Zou's confidence interval (Zou, 2007) available in the "cocor" R package (Diedenhofen & Musch, 2015) was used to test for significant differences between correlations both within and across samples.

Results

Sample characteristics

Table 1 summarizes the demographic and clinical characteristics of the samples used to address $Aim\ One$, comparing written versus oral tasks (n=144) and $Aim\ Two$, comparing written versus digital tasks (n=177). A slight majority of participants in both samples were female. A slight majority of participants in the $Aim\ One$ sample had pharmacoresistant focal epilepsy whereas a slight majority of participants in the $Aim\ Two$ sample had a new diagnosis of epilepsy.

Aim one: in-person written versus remote oral SDMT and TMT

Table 2 summarizes participant performance on the written and oral versions of the SDMT and TMT. For the SDMT, participants, on average, transcribed more symbols on the remotely administered oral rather than the in-person administered written format, as expected given the spoken versus written response. For the TMT, on average, as expected, participants completed the remotely administered oral TMT faster than the in-person administered written TMT.

Figure 1 includes scatterplots showing relationships between the written and oral SDMT (Figure 1-A), TMT-A (Figure 1-B), and TMT-B (Figure 1-C) raw scores. There was a strong relationship between the written and oral SDMT, with 59.4% shared variance, r (133) = .77, p < .001, 95% CI [.69, .83]. There was a weak relationship between the written and oral TMT-A, which shared

Table 1. Participant demographic and clinical characteristics

	Aim One: Written Vs. Oral $(n = 144)$			Aim Two: Written Vs. Digital ($n = 177$)			
	n (%)	Mean (SD)	Range	n (%)	Mean (SD)	Range	
Age		36.4 (12.3)	18.3 - 62.9		36.8 (13.9)	18.1 - 66.7	
Sex							
Male	65 (45.1)			78 (44.1)			
Female	79 (54.9)			99 (55.9)			
Years of education	130 ^b	13.3 (2.7)	7 – 20	177	13.1 (2.5)	7 – 20	
Referral category							
First unprovoked seizure	25 (17.4)			48 (27.1)			
New diagnosis of epilepsy	49 (34.0)			69 (39.0)			
Pharmacoresistant focal epilepsy	70 (48.6)			60 (33.9)			
Epilepsy type ^a							
Generalized	1 (0.8)			11 (8.5)			
Focal	98 (82.4)			105 (81.4)			
Unknown	20 (16.8)			13 (10.1)			
ToPF standard score	133 ^b	99.9 (13.2)	76 - 124	159 ^b	101.5 (13.6)	69 - 127	

Note: ToPF = Test of Premorbid Function.

Table 2. Summary statistics (Mean, SD, median, quartiles) of scores on the written and oral SDMT and TMT (Aim one)

		Written		Oral	
	n	Mean (SD)	Median (1 st Quartile – 3 rd Quartile)	Mean (SD)	Median (1 st Quartile – 3 rd Quartile)
SDMT total correct responses TMT-A	135	46.8 (12.2)	48.0 (37.5 – 55.0)	48.7 (12.0)	50 (42.5 – 55.0)
Completion time (seconds)	132	29.0 (11.7)	26.1 (21.0 - 33.7)	8.1 (2.6)	7.5 (6.6 – 9.3)
Number of errors TMT-B	127 ^a	0.1 (0.4)	0.0 (0.0 – 0.0)	_	· -
Completion time (seconds)	128	75.5 (36.0)	66.0 (48.8 - 94.2)	47.4 (30.7)	38.0 (23.1 - 59.2)
Number of errors	123 ^a (Written); 128 (Oral)	0.5 (0.9)	0.0 (0.0 - 1.0)	1.1 (1.2)	1.0 (0.0 – 2.0)

Note: SDMT = Symbol Digit Modalities Task; TMT = Trail Making Test.

 $^{^{}a}$ Written TMT error data for n=5 participants was missing.

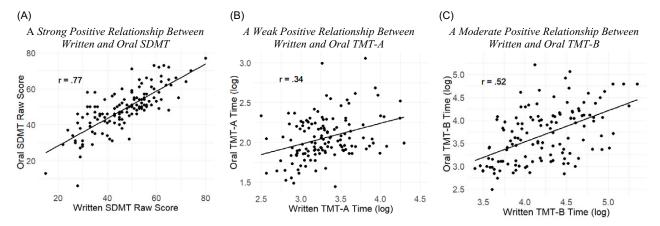


Figure 1. Written and oral SDMT are strongly correlated (A) while written and oral TMT are not (B, C).

only 11.4% variance, r (130) = .34, p < .001, 95% CI [.18, .48]. There was a moderate relationship between the written and oral TMT-B, with 26.8% shared variance, r (126) = .52, p < .001, 95% CI [.38, .63]. The correlations between the linear written and oral TMT were slightly weaker than those observed in the logarithmic transformed variables (TMT-A: r (130) = .30, p < .001, 95% CI [.14, .45]; TMT-B: r (127) = .41, p < .001, 95% CI [.26, .55]).

When administering the TMT (in any format), errors must be corrected in real time. Thus, when errors are made, completion time reflects both cognitive processing and error correction. In our sample the number of people who made errors (and the number of errors they made) was greater on the oral version of TMT-B (n = 70; 54.7%) than the written TMT-B (n = 45; 36.6%), and therefore a greater proportion of the completion time on this shorter task

^aSummary statistics are based on participants in the new diagnosis of epilepsy and pharmacoresistant focal epilepsy groups only. ^bTotal *n* lower than the total sample *n* due to missing data or non-complete or unreliable data.

Table 3. Written and oral TMT-B measure only partially overlapping cognitive constructs (Aim one)

	n	Mean (SD)	Written TMT-B ^a r	Oral TMT-B ^a <i>r</i>	r _{difference} - Oral Minus Written TMT-B [Zou's (2007) 95% CI] ^c
Oral TMT-B ^a	128		.52***	-	_
Written TMT-B ^a	128		-	.52***	-
Written SDMT	128		71***	46***	25 [.12, .39]
Written TMT-A ^a	128		.63***	.41***	.22 [.08, .36]
WAIS-IV Longest Digit Span Backward	122	4.9 (1.4)	40***	31***	-
Letter Fluency	120	36.0 (12.8)	45***	47***	-
CANTAB RVP z-score ^b	45	-0.32 (0.8)	57***	32*	_
CANTAB SWM z-scoreb	49	-0.1 (1.5)	46***	16	29 [55,03]
ToPF Standard Score	117	99.5 (13.0)	41***	37***	
Age	128	36.3 (12.4)	.08	08	-

Note: SDMT = Symbol Digit Modalities Task; TMT = Trail Making Test; WAIS-IV = Weschler Adult Intelligence Scale - Fourth Edition; RVP = Rapid Visual Processing; SWM = Spatial Working Memory; ToPF = Test of Premorbid Function.

Table 4. Summary statistics (Mean, SD, median, quartiles) of scores on the written SDMT and TMT and digital symbol decoding and TMT (Aim two)

		Written		Digital		
	n	Mean (SD)	Median (1 st Quartile – 3 rd Quartile)	Mean (SD)	Median (1 st Quartile – 3 rd Quartile)	
SDMT/Symbol Decoding total correct responses	128	47.5 (9.8)	47.0 (42.8 – 53.0)	50.5 (10.9)	50.0 (45.0 – 56.3)	
TMT-A						
Completion time (seconds)	162	26.5 (11.1)	24.5 (19.9 - 30.0)	31.9 (12.4)	28.8 (23.6 - 37.6)	
Number of errors	162	0.2 (0.4)	0.0 (0.0 - 0.0)	0.3 (0.8)	0.0 (0.0 - 0.0)	
TMT-B						
Completion time (seconds)	156	63.1 (26.1)	56.7 (45.4 - 74.6)	64.9 (26.2)	59.5 (47.6 - 76.4)	
Number of errors	156	0.4 (0.6)	0.0 (0.0 - 1.0)	0.9 (2.1)	0.0 (0.0 - 1.0)	

 $\it Note: SDMT = Symbol Digit Modalities Task; TMT = Trail Making Test.$

was spent correcting errors. The administrative 'time cost' of correcting these more frequent errors on the oral version likely contributes to the weaker relationship we observed between oral and written TMT-B compared to that reported elsewhere (Bastug et al., 2013; Mrazik et al., 2010; Ricker & Axelrod, 1994).

Relationships between written and oral TMT-B and other neuropsychological tasks and demographic variables

As noted in the Introduction, the different test properties of the oral and written TMT results in only partial overlap of the cognitive processes tapped by either task. This likely also contributes to the relatively weaker relationship we observed between written and oral TMT-B. Correlations of written and oral TMT-B with other related variables are reported in Table 3, with significant differences between the written and oral versions reported in bold.

The main difference is that the oral TMT-B is less strongly correlated with written TMT-A and written SDMT. This may reflect the fact that oral TMT-B is less dependent on processing speed. It also might reflect the fact that written TMT and written SDMT are both written tasks that, therefore, require visuomotor integration. Further, oral TMT-B is less dependent on working memory (CANTAB SWM and a trend in LDSB) than its written counterpart. There is also a tendency for sustained attention (CANTAB RVP) to be more strongly associated with written TMT-B performance than oral TMT-B, although not significantly so.

To summarize, while the written and oral forms of the SDMT appear comparable to one another, the written and oral forms of the TMT show a number of differences. In pursuit of a remotely administrable task that better replicates the original written TMT we next implemented and evaluated a videoconference-integrated, remotely administered, digital version of the TMT. We also compared a remotely administered digital SDMT analogous task, Symbol Decoding, with the original written form.

Aim two: in-person written SDMT and TMT versus remote digital symbol decoding and TMT

Table 4 summarizes participant performance on the in-person administered written SDMT and TMT and remotely administered digital Symbol Decoding and TMT. Participants transcribed more symbols on digital Symbol Decoding than written SDMT. Participants completed the written TMT-A slightly faster than they completed the digital TMT-A. The response times across the written and digital TMT-B were comparable. A similar percentage of participants made errors on the written (n = 43, 27.6%) and digital (n = 52, 33.3%) TMT-B.

Figure 2 includes scatterplots showing the relationships between the written SDMT and digital Symbol Decoding (Figure 2-A), written and digital TMT-A (Figure 2-B), and written and digital TMT-B (Figure 2-C) raw scores. There was a strong relationship between the written SDMT and digital Symbol Decoding, which shared 57.5% of variance, r(126) = .76, p < .001, 95% CI [.67, .82]. There was a moderate relationship between the

^aVariables have undergone a logarithmic transformation to ensure normality. ^bSpearman's correlations have been used due to the distributions being non-normal and not able to be transformed. ^cOnly significant results displayed, highlighted in bold (i.e., only results where Zou's (2007) 95% CI does not contain 0).

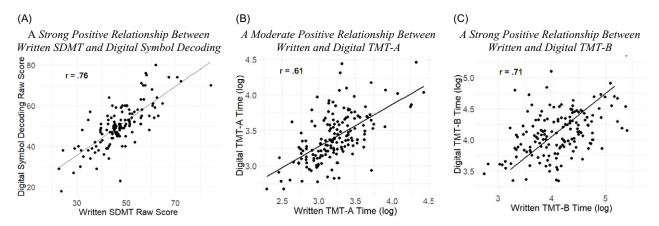


Figure 2. Written SDMT and digital symbol decoding and written and digital TMT are moderately-strongly correlated.

written and digital TMT-A, which shared 36.7% of variance, r (160) = .61, p < .001, 95% CI [.50, .69]. There was a strong relationship between the written and digital TMT-B, which shared 50.3% of variance, r (154) = .71, p < .001, 95% CI [.62, .78]. The correlations between the linear written and digital TMT were slightly weaker than those observed in the logarithmic transformed variables (TMT-A: r (160) = .57, p < .001, 95% CI [.45, .66]; TMT-B: r (154) = .69, p < .001, 95% CI [.60, .76]).

There was no difference in the correlation between the oral and written SDMT (.77) and the written SDMT and digital Symbol Decoding (.76), $r_{\rm difference} = .01$, Zou's (2007) CI [-0.09, 0.12]. The correlation between written TMT-B and digital TMT-B (r = .71) was significantly greater than the correlation between the written TMT-B and oral TMT-B (r = .52), $r_{\rm difference} = -.19$, Zou's (2007) CI [-0.35, -0.05]. This was also the case for TMT-A, $r_{\rm difference} = 0.27$, Zou's (2007) CI [0.09, 0.45]. These comparisons highlight that the digital TMT more closely captures the properties of the original written TMT than does the oral TMT.

Discussion

Here we assess the suitability of TeleNP-compatible versions of the SDMT and TMT (i.e., remotely administered oral and digital versions), by evaluating their comparability to the sensitive and frequently used written SDMT and TMT in people with seizure disorders. Our data show that remotely administered versions of the SDMT, oral or digital, are comparable to in-person administrations of the original written SDMT task. Conversely, for the TMT, using the oral form for remote administration fundamentally changes the nature of the task. The novel remotely administered digital TMT described here much more closely reproduces the original written form of the task.

SDMT and digital symbol decoding

As hypothesized, both the remotely administered oral SDMT and digital Symbol Decoding were strongly correlated with the inperson administered written SDMT. These correlations were broadly in keeping with the established test–retest reliability of .80 for the written SDMT itself (Smith, 1991). This is consistent with research in other clinical samples that similarly supports the administration of the oral SDMT via TeleNP (e.g., dementia: Brown et al., 2024; stroke: Chapman et al., 2021; MS: Eilam-Stock et al., 2021). Our study indicates that both the remotely administered oral SDMT and digital Symbol Decoding are possible

alternatives to the in-person written SDMT. This is perhaps unsurprising given that the nature of the three tasks does not fundamentally vary across versions, the only difference being that the written SDMT requires a written response while the oral SDMT and digital Symbol Decoding require a spoken response. The implications of this are widespread and go beyond accommodating for infection control measures as a result of COVID-19 to include such benefits as potential use with groups with upper limb mobility restrictions (e.g., those with stroke, spinal injury, demyelinating diseases, or movement disorders), overcoming barriers to service access (e.g., driving restrictions, a scarcity of clinicians in non-metropolitan areas), and allowing for geographically dispersed recruitment while maintaining control and standardization of data collection, as is being done in the AEP.

In a TeleNP setting, the novel digital Symbol Decoding task described here has benefits over the oral SDMT. Digital Symbol Decoding is highly standardized as the instructions are built into the software and therefore easily adhered to, and the timing is automated. Responses are recorded by the examiner directly into the software and the task is automatically scored. This saves the examiner time and minimizes (although does not eliminate) the scope for human error in scoring. Digital Symbol Decoding also provides a moment-by-moment record of participant responses across time, which means it has the potential to provide examiners with additional clinically informative metrics to characterize performance (e.g., variations in performance throughout the trial). Further research, however, is required to characterize and evaluate such metrics.

TMT

As hypothesized, the written and oral TMT-A shared only a weak relationship. This finding is consistent with previous suggestions that the automatic and overlearned nature of oral TMT-A (i.e., counting from 1 – 25) reduces its sensitivity to deficits in processing speed and attention relative to the written form (Bastug et al., 2013; Jaywant et al., 2018; Mrazik et al., 2010). Despite this, when using the oral TMT (e.g., during telephone-based cognitive screening; e.g., Pugh, Vaughan, Jackson, Ponsford, & Tailby, 2024) it is likely still important to administer oral TMT-A to establish a response set that the examinee must subsequently modify in order to complete oral TMT-B. In line with our hypothesis, the written and oral TMT-B shared only a moderate relationship in our study. This seems to reflect both a) that the nature of the oral and written TMT-B differs sufficiently such that they measure only partially

overlapping cognitive constructs, and b) that error correction likely impacts completion time disproportionately across task versions. As shown in Table 3, oral TMT-B seems to rely less on processing speed and attention, particularly sustained attention, due to its shorter completion times relative to written TMT-B. The written task also has a visual search element that is not present in its oral counterpart, making the written form more dependent upon spatial working memory and related visual search strategies. The written task also has a motor component not present in the oral task. Overall, it appears the written TMT-B is more multidetermined. Given that the written and oral TMT-B share less than a third of their variance, clinicians should be cautious about drawing on the extant literature on written TMT-B when interpreting performance on the oral TMT-B (Bastug et al., 2013). We also suspect that the oral form will exhibit lower reliability, especially in clinical cohorts, given the greater proportional impact of errors on completion time, compounded by the overall shorter completion time, though we do not have data to address this question directly. Even so, this measure may be most valuable as a screening measure whereby obvious failure prompts further assessment. It is worth noting that errors may be more prevalent in our clinical sample than in previously evaluated healthy samples, which may explain the relatively weaker correlation between written and oral TMT-B here, compared to that reported elsewhere (Bastug et al., 2013; Mrazik et al., 2010; Ricker & Axelrod, 1994). While the oral form of TMT certainly remains useful in certain contexts (e.g., telephone-based screening; Pugh et al., 2024, the visually or upper limb mobility impaired, possibly as a purer measure of attentional switching), where possible, it appears a visual TMT is preferable given the superior psychometric properties of this task version and the broader understanding of its nature.

To overcome these shortcomings of the oral version, we developed and tested a remotely administrable, videoconferenceintegrated digital version of the TMT. We observed a moderate relationship between the written and digital TMT-A. The written and digital TMT-B were strongly correlated. Importantly, the correlation between written and digital TMT-B (0.71) is comparable to the test-retest reliability of the written TMT-B itself (between .60 and .90; Strauss et al., 2006), the upper bound against which a written-digital comparison can be judged. The slightly weaker correlation for TMT-A compared to TMT-B likely reflects the slightly weaker reliability of TMT-A itself (Strauss et al., 2006). Overall, it appears that the digital TMT more closely resembles the well-understood written TMT and is, therefore, a better alternative in a TeleNP context. Version-appropriate norms for all measures administered via TeleNP are clearly important, given the observed mean differences in performance between versions. We are currently collecting control data for all tasks in our custom TeleNP software to create appropriate age and demographically adjusted norms.

The digital TMT has several advantages over the oral TMT, beyond better approximating the written TMT. These include the increased standardization and automation in administration and scoring that the software facilitates and the potential to further characterize performance with additional clinically useful metrics. For example, further work may evaluate (a) the clinical utility of evaluating different error types (e.g., sequencing errors versus switching errors), (b) whether slowed completion times can be

attributed to different response patterns (e.g., general slowness, isolated visual search difficulties), and (c) search times for stimuli in each visual field in the context of visual field deficits such as hemispatial neglect. Importantly, while we have used digital tasks here to facilitate remote testing, the aforementioned benefits could also be obtained using in-person administration of these digital tasks.

Limitations

In our study, different groups of examiners administered the written versus the oral versions of each measure and the written versus the digital versions of each measure. We sought to mitigate this potential source of bias via extensive training and the use of standardized instructions. We acknowledge that the retrospective evaluations of performance validity on the written tasks were less ideal than the real-time evaluations made for oral and digital tasks. The order in which the written and oral and written and digital tasks versions were administered, and the test form versions in the case of the SDMT, were not counterbalanced. It also would have been beneficial to include an additional measure of attentional switching to better characterize the content validity of the TMT. Further, we acknowledge the use of a relatively mildly impaired clinical sample here, and these conclusions should be limited to this group. Future studies can address these issues, along with issues of test-retest reliability which were not explored here.

Summary and conclusion

Both the remotely administered oral SDMT and digital Symbol Decoding are highly correlated with the original in-person administered written SDMT in a cohort of individuals with seizure disorders. Any of these task variants can serve as a useful marker of processing speed and attention. Oral TMT-A is automatic and overlearned and as such does not correlate well with written TMT-A. Written and oral TMT-B measure only partially overlapping cognitive constructs. The written and digital TMT, however, were more highly correlated and appear to measure similar constructs. The digital TMT is more suitable in a TeleNP setting. Digital tasks offer additional benefits over their oral (and written) forms. Appropriate normative data will enhance their usefulness. This study highlights the benefits of capitalizing on the TeleNP modality, not only in terms of increasing the reach and accessibility of neuropsychology, but also by using technology to facilitate task administration and scoring and to automatically capture otherwise difficult-to-obtain task metadata. These latter benefits apply not only to remote TeleNP assessments but to inperson administration using these same digital assessment procedures.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1355617725101124.

Acknowledgements. The authors would like to thank the participants for volunteering their time, the referring neurologists, and the following AEP personnel: Emma Beildeck, Seiko Bhungane, Elliot Brooker, Marty Bryant, Alana Collins, Christina Fairley, Chathurini Fernando, Jocelyn Halim, Elise Honey, Rachel Hughes, Molly Ireland, Loren Lindenmayer, Evie Muller,

Donna Parker, Isobella Peterson, Remy Pugh, Laura Stevens, Carmen Zheng.

Funding statement. The Australian Epilepsy Project received funding from the Australian Government under the Medical Research Future Fund (Frontier Health and Medical Research Program - Grant Numbers MRFF75908 and RFRHPSI000008) and the Victoria State Government (Victorian-led Frontier Health and Medical Research Program). DFA acknowledges fellowship funding from the National Imaging Facility. The Florey Institute of Neuroscience and Mental Health acknowledges the strong support from the Victorian Government and in particular the funding from the Operational Infrastructure Support Grant.

Competing interests. The authors declare that there are no conflicts of interest.

References

- Akbar, N., Honarmand, K., Kou, N., & Feinstein, A. (2011). Validity of a computerized version of the symbol digit modalities test in multiple sclerosis. *Journal of Neurology*, 258, 373–379.
- Bastug, G., Ozel-Kizil, E. T., Sakarya, A., Altintas, O., Kirici, S., & Altunoz, U. (2013). Oral trail making task as a discriminative tool for different levels of cognitive impairment and normal aging. Archives of Clinical Neuropsychology, 28(5), 411–417.
- Baykara, E., Kuhn, C., Linz, N., Tröger, J., & Karbach, J. (2022). Validation of a digital, tablet-based version of the trail making test in the Δelta platform. European Journal of Neuroscience, 55(2), 461–467.
- Bigi, S., Marrie, R. A., Till, C., Yeh, E. A., Akbar, N., Feinstein, A., & Banwell, B. L. (2017). The computer-based symbol digit modalities test: Establishing age-expected performance in healthy controls and evaluation of pediatric MS patients. *Neurological Sciences*, 38(4), 635–642.
- Brown, A. D., Kelso, W., Eratne, D., Loi, S. M., Farrand, S., Summerell, P., Neath, J., M., Walterfang, Velakoulis, D., & Stolwyk, R. J. (2024). Investigating equivalence of in-person and telehealth-based neuro-psychological assessment performance for individuals being investigated for younger onset dementia. Archives of Clinical Neuropsychology, 39(5), 594–607.
- Cambridge Cognition. (2019). [Computer software].
- Chapman, J., Gardner, B., Ponsford, J., Cadilhac, D., & Stolwyk, R. (2021). Comparing performance across in-person and videoconference-based administrations of common neuropsychological measures in communitybased survivors of stroke. *Journal of the International Neuropsychological Society*, 27(7), 697–710.
- Chapman, J., Ponsford, J., Bagot, K. L., Cadilhac, D., Gardner, B., & Stolwyk, R. (2020). The use of videoconferencing in clinical neuropsychology practice: A mixed methods evalutation of neuropsychologists' experiences and views. Australian Psychologist, 55(6), 618–633.
- Dahmen, J., Cook, D., Fellows, R., & Schmitter-Edgecombe, M. (2017). An analysis of a digital variant of the trail making test using machine learning techniques. *Technology and Health Care*, 25(2), 251–264.
- Diedenhofen, B., & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations (Version 1.1-4). https://doi.org/10. 1371/journal.pone.0121945.
- Eilam-Stock, T., Shaw, M. T., Sherman, K., Krupp, L. B., & Charvet, L. E. (2021).
 Remote administration of the symbol digit modalities test to individuals with multiple sclerosis is reliable: A short report. *Multiple Sclerosis Journal Experimental, Translational and Clinical*, 7(1), 1–3.
- Fellows, R. P., Dahmen, J., Cook, D., & Schmitter-Edgecombe, M. (2017).
 Multicomponent analysis of a digital trail making test. The Clinical Neuropsychologist, 31(1), 154–167.

- Forn, C., Belloch, V., Bustamante, J. C., Garbin, G., Parcet-Ibars, M. D., Sanjuan, A., Ventura, N., & Ávila, César (2009). A symbol digit modalities test version suitable for functional MRI studies. *Neuroscience Letters*, 456(1), 11–14.
- Grigsby, J., & Kaye, K. (1995). Alphanumeric sequencing and cognitive impairment among elderly persons. *Perceptual and Motor Skills*, 80(3), 732–734.
- Hardy, D. J., Castellon, S. A., & Hinkin, C. H. (2021). Incidental learning and memory deficits on a computerized symbol-digit modalities test in adults with HIV/AIDS. *Journal of the International Neuropsychological Society*, 27(4), 389–395
- Hinton-Bayre, A., & Geffen, G. (2005). Comparability, reliability, and practice effects on alternate forms of the digit symbol substitution and symbol digit modalities tests. *Psychological Assessment*, 17(2), 237–241.
- Jaywant, A., Barredo, J., Ahern, D. C., & Resnik, L. (2018). Neuropsychological assessment without upper limb involvement: A systematic review of oral versions of the trail making test and symbol-digit modalities test. Neuropsychological Rehabilitation, 28(7), 1055–1077.
- Kitaigorodsky, M., Loewenstein, D., Curiel Cid, R., Crocco, E., Gorman, K., & González-Jiménez, C. (2021). A teleneuropsychology protocol for the cognitive assessment of older adults during COVID-19. Frontiers in Psychology, 12, 1–6.
- Lunardini, F., Luperto, M., Daniele, K., Basilico, N., Damanti, S., Abbate, C., Mari, D., Cesari, M., Ferrante, S., Borghese, N. A. (2019). Validity of digital trail making test and bells test in elderlies. 2019 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI), 4.
- Lutz, M. T., & Helmstaedter, C. (2005). EpiTrack: Tracking cognitive side effects of medication on attention and executive functions in patients with epilepsy. *Epilepsy & Behavior*, 7(4), 708–714.
- Marvel, C. L., & Paradiso, S. (2004). Cognitive and neurological impairment in mood disorders. The Psychiatric Clinics of North America, 27(1), 19.
- Mrazik, M., Millis, S., & Drane, D. L. (2010). The oral trail making test: Effects of age and concurrent validity. Archives of Clinical Neuropsychology, 25(3), 236–243.
- Nys, G. M. S., van Zandvoort, M. J. E., de Kort, P. L. M., Jansen, B. P. W., de Haan, E. H. F., & Kappelle, L. J. (2007). Cognitive disorders in acute stroke: Prevalence and clinical determinants. *Cerebrovascular Diseases*, 23(5-6), 408–416.
- Oreja-Guevara, C., Ayuso Blanco, T., Brieva Ruiz, L., Hernández Pérez, M.Á., Meca-Lallana, V., & Ramió-Torrentà, L. (2019). Cognitive dysfunctions and assessments in multiple sclerosis. *Frontiers in Neurology*, 10, 581.
- Park, S. Y., & Schott, N. (2022). The trail-making-test: Comparison between paper-and-pencil and computerized versions in young and healthy older adults. Applied Neuropsychology: Adult, 29(5), 1208–1220.
- Ponsford, J., Bayley, M., Wiseman-Hakes, C., Togher, L., Velikonja, D., McIntyre, A., Janzen, S., & Tate, R. (2014). INCOG recommendations for management of cognition following traumatic brain injury, part II. *The Journal of Head Trauma Rehabilitation*, 29, 321–337.
- Ponsford, J., & Kinsella, G. (1992). Attentional deficits following closed-head injury. Journal of Clinical and Experimental Neuropsychology, 14(5), 822–838.
- Ponsford, J., Sloan, S., & Snow, P. (2013). Traumatic Brain Injury: Rehabilitation for Everyday Adaptive Living (2nd edn.). Psychology Press. https://doi.org/10.4324/9780203082805
- Pugh, R., Vaughan, D. N., Jackson, G. D., Ponsford, J., & Tailby, C. (2024). Cognitive and psychological dysfunction is present after a first seizure, prior to epilepsy diagnosis and treatment at a first seizure clinic. *Epilepsia Open*, 9(2), 717–726.
- Ricker, J. H., & Axelrod, B. N. (1994). Analysis of an oral paradigm for the trail making test. Assessment, 1(1), 47–51.
- RStudio Team. (2021). RStudio (Version 1.4.1106) [Computer software]
- Sandroff, B. M., Pilutti, L. A., Dlugonski, D., & Motl, R. W. (2013). Physical activity and information processing speed in persons with multiple sclerosis: A prospective study. *Mental Health and Physical Activity*, 6(3), 205–211.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 3(4), 591–611.

- Smith, A. (1991). Symbol Digit Modalities Test. Western Psychological Services.Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). A Compendium of Neuropsychological Tests: Administration, Norms and Commentary (3rd edn.) Oxford University Press.
- Tailby, C., Chapman, J., Pugh, R., Holth Skogan, A., Helmstaedter, C., & Jackson, G. D. (2024). Applications of teleneuropsychology to the screening and monitoring of epilepsy. Seizure: European Journal of Epilepsy, 128, 54–58.
- Tailby, C., Collins, A. J., Vaughan, D. N., Abbott, D. F., O'Shea, M., Helmstaedter, C., & Jackson, G. D. (2020). Teleneuropsychology in the time of COVID-19: The experience of the Australian epilepsy project. Seizure, 83, 89–97.
- van Rijckevorsel, K. (2006). Cognitive problems related to epilepsy syndromes, especially malignant epilepsies. *Seizure*, 15(4), 227–234.
- Wechsler, D. (2008). Wechsler Adult Intelligence Scale—Fourth Edition, Australian and New Zealand Adapted Edition. Pearson Clinical and Talent Assessment.
- Wechsler, D. (2009). Advanced Clinical Solutions for WAIS-IV and WMS-IV. Pearson Clinical and Talent Assessment.
- Zane, K. L., Thaler, N. S., Reilly, S. E., Mahoney, J. J., & Scarisbrick, D. M. (2021). Neuropsychologists' practice adjustments: The impact of COVID-19. *The Clinical Neuropsychologist*, 35(3), 490–517.
- Zoom Video Communications, Inc (2021). Zoom [Computer software].
- Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. Psychological Methods, 12(4), 399–413.