Comparing Performance Across In-person and Videoconference-Based Administrations of Common Neuropsychological Measures in Community-Based Survivors of Stroke

Jodie E. Chapman[1] , Betina Gardner[1,2] , Jennie Ponsford[1,3] , Dominique A. Cadilhac[4,5] and
Renerus J. Stolwyk[1,3,*]

[1]School of Psychological Sciences and Turner Institute for Brain and Mental Health, Monash University, Melbourne, Victoria, Australia
[2]Department of Psychiatry, Monash University, Melbourne, Victoria, Australia
[3]Monash Epworth Rehabilitation Research Centre, Epworth HealthCare, Melbourne, Victoria, Australia
[4]Stroke & Ageing Research, Department of Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, Victoria, Australia
[5]Stroke Division, Florey Institute of Neuroscience and Mental Health, The University of Melbourne, Heidelberg, Victoria, Australia

Abstract

Objective: Neuropsychological assessment *via* videoconference could assist in bridging service access gaps due to geographical, mobility, or infection control barriers. We aimed to compare performances on neuropsychological measures across in-person and videoconference-based administrations in community-based survivors of stroke. **Method:** Participants were recruited through a stroke-specific database and community advertising. Stroke survivors were eligible if they had no upcoming neuropsychological assessment, concurrent neurological and/or major psychiatric diagnoses, and/or sensory, motor, or language impairment that would preclude standardised assessment. Thirteen neuropsychological measures were administered in-person and *via* videoconference in a randomised crossover design (2-week interval). Videoconference calls were established between two laptop computers, facilitated by Zoom. Repeated-measures *t* tests, intraclass correlation coefficients (ICCs), and Bland–Altman plots were used to compare performance across conditions. **Results:** Forty-eight participants (26 men; M_{age} = 64.6, SD = 10.1; $M_{time\ since\ stroke}$ = 5.2 years, SD = 4.0) completed both sessions on average 15.8 (SD = 9.7) days apart. For most measures, the participants did not perform systematically better in a particular condition, indicating agreement between administration methods. However, on the Hopkins Verbal Learning Test – Revised, participants performed poorer in the videoconference condition (Total Recall $M_{difference}$ = −2.11). ICC estimates ranged from .40 to .96 across measures. **Conclusions:** This study provides preliminary evidence that in-person and videoconference assessment result in comparable scores for most neuropsychological tests evaluated in mildly impaired community-based survivors of stroke. This preliminary evidence supports teleneuropsychological assessment to address service gaps in stroke rehabilitation; however, further research is needed in more diverse stroke samples.

Keywords: Cerebrovascular disorders, Cognition, Comparative effectiveness, Neuropsychology, Telehealth, Teleneuropsychology

INTRODUCTION

In Australia and internationally, neuropsychologists are primarily located in metropolitan areas (Australian Institute of Health and Welfare, 2018; Janzen & Guger, 2016; Psychology Board of Australia, 2018; Sweet, Benson, Nelson, & Moberg, 2015). *Teleneuropsychology,* defined

as the provision of neuropsychological services *via* telecommunication technologies, particularly videoconference (Cullum & Grosch, 2013), has the potential to expand the reach of neuropsychological services to those in underserviced areas. Beyond this, videoconference-based consultations could increase access to neuropsychology for those with mobility restrictions and could increase engagement and representation in research. More recently, the potential of teleneuropsychology to accommodate infection control measures such as those put in place as a result of COVID-19 has also been realised. Emerging evidence

*Correspondence and reprint requests to: Renerus J. Stolwyk, School of Psychological Sciences and Turner Institute for Brain and Mental Health, 18 Innovation Walk, Monash University, Melbourne, Victoria 3800, Australia. E-mail: rene.stolwyk@monash.edu

supports the use of videoconferencing for various aspects of neuropsychological practice including taking a clinical history (e.g., Martin-Khan et al., 2012; Martin-Khan, Varghese, Wootton, & Gray, 2007; Schopp, Johnstone, & Merveille, 2000) and providing cognitive interventions (e.g., Burton & O'Connell, 2018; Lawson et al., 2020).

Neuropsychological assessment (in particular, test administration) remains at the core of neuropsychological practice in most settings (Ponsford, 2016). This is also the area of practice in which neuropsychologists are most apprehensive and least confident about the use of videoconference (Chapman et al., 2020). Researchers have compared in-person and videoconference-based neuropsychological test administration in several populations including healthy individuals (Cullum, Hynan, Grosch, Parikh, & Weiner, 2014; Hildebrand, Chow, Williams, Nelson, & Wass, 2004; Jacobsen, Sprenger, Andersson, & Krogstad, 2003; Rebchuck et al., 2019; Wadsworth et al., 2016, 2018), and those with mild cognitive impairment (MCI; Cullum et al., 2014; Wadsworth et al., 2016, 2018), Alzheimer's disease (AD; Cullum et al., 2014; Wadsworth et al., 2018), unspecified dementia (Wadsworth et al., 2016), early psychosis (Stain et al., 2011), a history of alcohol abuse (Kirkwood, Peck, & Bennie, 2000), and developmental disorders (Temple, Drummond, Valiquette, & Jozsvai, 2010). Results from this body of research are broadly promising (for a review, see Brearly et al., 2017; Marra, Hamlet, Bauer, & Bowers, 2020). The authors of a recent meta-analysis demonstrated that for non-timed tests that allow for repetition, videoconference scores were one-tenth of a standard deviation below in-person scores (Brearly et al., 2017). In contrast, in-person and videoconference scores were equivalent for verbally mediated, timed tests that proscribe repetition (e.g., list learning tasks; Brearly et al., 2017). Client evaluations of acceptability have also been broadly positive (Hildebrand et al., 2004; Hodge et al., 2019; Kirkwood et al., 2000; Jacobsen et al., 2003; Parikh et al., 2013; Stain et al., 2011). However, to date, no research has compared performance scores between in-person and videoconference-based administrations of neuropsychological measures, nor examined patient acceptability of this method of assessment, in stroke survivors. It is important to investigate the suitability of videoconference-based neuropsychological test administration in stroke survivors. It cannot be assumed that the results of the above studies are generalisable to a stroke population who present with a diverse range of cognitive impairments. Further, many of the neuropsychological measures commonly used in the assessment of stroke survivors have not been previously evaluated.

Cognitive impairment is very common following stroke (Lesniak, Bak, Czepiel, Seniow, & Czlonkowska, 2008). While estimates of the rate of post-stroke cognitive impairment vary from approximately 30% to 90% depending on various factors such as the method of recruitment used in the study (e.g., hospital or community recruitment) and the criteria used to define cognitive impairment, the most reliable estimates suggest that cognitive impairment occurs in over 70% of stroke survivors (Lesniak et al., 2008). It is important to assess post-stroke cognition in order to plan and guide effective multidisciplinary rehabilitation and because of its prognostic value with regard to long-term outcomes (Nys et al., 2006; Saxena, Ng, Koh, Yong, & Fong, 2007; Wagle et al., 2011). Indeed, the authors of Australian clinical guidelines recommend (a) all stroke survivors should undergo cognitive screening and (b) where screening indicates likely cognitive impairment, a full neuropsychological evaluation should be undertaken (Stroke Foundation, 2017). However, reflecting the above-reported disparate distribution of neuropsychologists, these recommendations are not currently being met (Stroke Foundation, 2018, 2019). Based on the most recent Australian audit data, only 30% of acute stroke services have access to neuropsychology, which includes 46% of metropolitan services and just 10% of inner regional services and 13% of outer regional services (Stroke Foundation, 2019). Further, only 41% of rehabilitation services have access to neuropsychology (Stroke Foundation, 2018). Clearly, increasing access to neuropsychological services, by conducting neuropsychological assessments *via* videoconference, could benefit stroke survivors (and their carers) by increasing the identification, characterisation, and management of cognitive impairment post-stroke.

Aims and Hypotheses

Our primary aim was to compare performance across in-person and videoconference-based administrations of common neuropsychological tasks in community-based survivors of stroke. On the basis of previous research, we hypothesised that performance in in-person and videoconference conditions would be comparable for all tests. The secondary aim was to evaluate the level of acceptability of videoconference-based neuropsychological assessment to participants. On the basis of previous research, we hypothesised that participants would show a high degree of acceptability of videoconference-based neuropsychological assessment.

METHOD

The design, procedure, and sample for this study have previously been published in Chapman et al. (2019), which presents results comparing in-person and videoconference-based administrations of the Montreal Cognitive Assessment (MoCA). The study design is provided briefly below, with greater detail on aspects of the study not provided in the initial paper published from this research.

Design

In-person and videoconference sessions were conducted in a counterbalanced order (randomised crossover design). We aimed for an interval of 2 weeks between sessions to balance the impact of practice effects and natural changes/fluctuations in cognition. For the Hopkins Verbal Learning
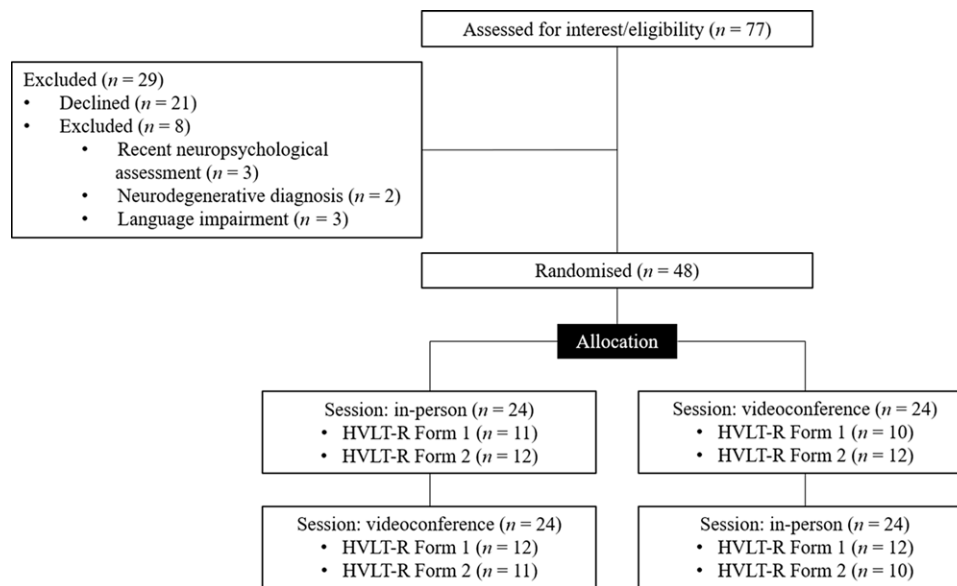
Fig. 1. Participant recruitment, progression, and counterbalancing. HVLT-R = Hopkins Verbal Learning Test – Revised.

Test – Revised (HVLT-R), alternate forms (Forms 1 and 2; Brandt & Benedict, 2001) were used. Form versions were counterbalanced on an opposite schedule to condition order. For example, half of the participants who completed the in-person session first were administered Form 1 in this session and Form 2 in the videoconference session while the other half were administered the forms in the opposite order across sessions (see Figure 1).

Participants

Participants were community-dwelling survivors of stroke recruited through community advertising and a stroke-specific database of former research participants. Participants were required to have a confirmed diagnosis of stroke and to be older than 18 years, proficient in English, and at least 3 months post-stroke, to avoid the most rapid period of spontaneous recovery (Skilbeck, Wade, Hewer, & Wood, 1983). People were excluded if they had (a) a recent or upcoming neuropsychological assessment for clinical purposes, (b) a concurrent neurological diagnosis other than stroke and/or a major psychiatric diagnosis(es), and/or (c) any sensory, motor or language impairment that would significantly preclude the standardised unadapted administration of tests. Sensory, motor, and language impairments were screened for in the initial contact with the participant and/or their carer. Where participants had language impairment but were still able to understand the study and provide consent, they were included and their capacity to complete each measure was assessed through the administration of the initial instructions and sample items. Where language impairment was likely to preclude valid administration, this measure was excluded. There were no exclusion criteria regarding access to technology, as we provided the required technology. Data were collected between November 2016 and February 2019 in Melbourne, Australia and surrounding regional areas.

Measures

Participant characteristics

Demographic data were obtained using a verbally administered questionnaire. Stroke information (e.g., mechanism and location of stroke) was sought from the participant's medical records (from their acute treating hospital or general practitioner) with their written consent. The following measures were administered to characterise the sample.

Cognitive screen. The MoCA is a 30-point cognitive screening measure (Nasreddine et al., 2005). Items assess visuospatial/executive function, naming, attention, language, abstraction, delayed recall, and orientation (Nasreddine et al., 2005). Conventionally, scores of ≤25 are considered indicative of likely cognitive impairment (Nasreddine et al., 2005). However, some authors suggest that a cutoff of ≤24 is more sensitive and specific in chronic stroke samples (Pendlebury, Mariz, Bull, Mehta, & Rothwell, 2012, 2013). Research supports the psychometric properties of the MoCA for use following stroke (Burton & Tyson, 2015). To allow for comparison of MoCA performance across in-person and videoconference administrations (presented in Chapman et al., 2019), alternate versions were administered in both conditions (in the same counterbalanced design as the neuropsychological measures). As such, where MoCA scores are reported and used in this paper, an average of these two administrations has been used.

Computer proficiency. The Computer Proficiency Questionnaire (CPQ) is a 33-item self-report measure of frequency and ease of computer use across six categories (e.g.,

Computer Basics; Boot et al., 2015). Responses to items (e.g., "I can: Turn a computer on and off") are provided on a 5-point scale ranging from 1 (*never tried*) to 5 (*very easily*); average scores in each category are summed to obtain a total score between 5 (low computer proficiency) and 30 (high computer proficiency; Boot et al., 2015). This measure has sound psychometric properties (Boot et al., 2015).

Mood. The Hospital Anxiety and Depression Scale (HADS) is a 14-item self-report measure that assesses symptoms of anxiety (HADS-A) and depression (HADS-D; Zigmond & Snaith, 1983). Items (e.g., HADS-A: "I feel tense or 'wound-up'") are answered on a scale from 0 (indicating the least frequent occurrence; e.g., *not at all*) to 3 (indicating the most frequent occurrence; e.g., *most of the time*; Zigmond & Snaith, 1983). Subscale scores are summed to reflect either normal (0–7), mild (8–10), moderate (11–14), or severe (15–21) symptomatology (Zigmond & Snaith, 1983). This measure has sound psychometric properties (Bjelland, Dahl, Haug, & Neckelmann, 2002; Zigmond & Snaith, 1983).

Neuropsychological measures

We evaluated neuropsychological measures across a range of cognitive domains, specifically, premorbid ability, attention and processing speed, language, visuospatial function, visual and verbal learning and memory, and executive function. We evaluated measures commonly and frequently used in the assessment of stroke survivors, with measure selection guided by the *common data elements* developed by the National Institute of Neurological Disorders and Stroke and Canadian Stroke Network (Hachinski et al., 2006). We refined the 60-min battery defined therein, based on consultation with experts in stroke rehabilitation (namely, Betina Gardner, Jennie Ponsford, and Renerus Stolwyk). In refining this battery, the oral Symbol Digit Modalities Test (SDMT; Smith, 1973) substituted the Wechsler Adult Intelligence Scale – Third Edition (WAIS-III; Wechsler, 1997) Digit Symbol Coding subtest to accommodate stroke survivors with upper limb impairments. Additional measures (e.g., Stroop Test [Victoria Modification]) were also added to expand the assessment of some domains. The neuropsychological battery is shown in Table 1. Although some measures assess multiple domains, they have been reported in only one domain.

Acceptability of administration methods

We used a 14-item self-report survey of acceptability, modified from the measure developed by Parikh et al. (2013), including two questions about participants' experience in the in-person condition, four questions about the videoconference condition, and eight questions comparing participants' experience across conditions (see Supplementary Material).

Videoconference setup

Videoconference calls were established between two laptop computers and facilitated using a cloud-based software, Zoom (Copyright © 2019 Zoom Video Communications, Inc., San Jose, California, USA). Videoconference calls had an established bandwidth of 384 kb/s, which is sufficient for a one-to-one video call (Bartlett & Wetzel, 2010) and has been deemed appropriate in similar studies (e.g., Jacobsen et al., 2003). We utilised the integrated webcam of each laptop directed to obtain a portrait view of the researcher/participant. An additional USB-connected webcam on the participant's laptop (located next to the integrated webcam) was directed so the researcher could observe the participant's work station, and therefore, their performance of tasks. Cameras were switchable by the participant using a two-key command; they were trained to do this at the beginning of the videoconference session. The integrated webcam was used for verbal tasks, and the USB webcam was used where participants were required to interact with stimulus materials. Images depicting this setup are provided in the Supplementary Material for Chapman et al. (2019).

Procedure

This research was completed in accordance with the Helsinki Declaration. Ethics approval was provided by the Monash University Human Research Ethics Committee (CF16/130 – 2016000056). We obtained written informed consent and demographic data in the first session.

Both sessions were conducted by the same researcher (Jodie Chapman), at the same time of day (where possible), at the same location (participant's home, university, or community location), and in a quiet, distraction-free room(s). In both conditions, neuropsychological tasks and the MoCA were administered in a predefined order, which minimised cross-task interference. All tests were administered in accordance with standardised administration instructions set out in test manuals. Modifications to standardised procedures in the videoconference condition are provided in Table 1. For three tasks (i.e., BNT, WAIS-IV Block Design, and WMS-IV VR), a research assistant was present to physically engage with test stimuli (e.g., turn pages) in the videoconference condition only. They were not required to provide instructions, time responses, or otherwise assist with task administration. There were several research assistants in this study who were all trained using a consistent protocol, although each participant was only exposed to one. In accordance with standard practice, task responses were recorded on response forms and scored after each session. Participants completed the CPQ, HADS, and acceptability survey at the end of the second session.

Data Analyses

Data analyses were conducted using IBM SPSS Statistics for Windows, Version 23.0. Scores compared for each measure

Table 1. Neuropsychological measures including administration modifications utilised in the videoconference condition and scores for comparison

Measure	Administration modification for videoconference condition	Scores for comparison
Premorbid ability		
Test of Premorbid Function (TOPF; Wechsler, 2009a)	*Word card* in an envelope at the participant's location; standard instructions provided	• Number of items correct
Attention and processing speed		
SDMT (Smith, 1973)	*Response form* in an envelope at the participant's location; standard instructions provided with pointing omitted	• Number of correct responses
WAIS – Fourth Edition (WAIS-IV) – Digit Span (Wechsler, 2008)	None required	• Number of items correct for: o Digit Span Total o Digit Span Forward (DSF) o Digit Span Backward (DSB) o Digit Span Sequencing (DSS)
Language		
Boston Naming Test – 2nd Edition (BNT; Kaplan, Goodglass, & Weintraub, 2001)	*Stimulus book* at the participant's location; research assistant changed pages in accordance with the examiner's instruction	• Number of items correct
Semantic Fluency (Animals; Strauss, Sherman, & Spreen, 2006)	None required	• Number of correct responses
Visuospatial function		
WAIS-IV Block Design (Wechsler, 2008)	*Stimulus book 1* and *blocks* at the participant's location; research assistant handled blocks and changed pages in accordance with the examiner's instruction; standard instructions provided with pointing omitted	• Total raw score
Learning and memory		
HVLT-R (Brandt & Benedict, 2001)	None required	• Number of correct responses for: o Total Recall (i.e., Trials 1–3) o Delayed Recall (i.e., Trial 4) • Discrimination Index (i.e., correct responses – false-positive errors on recognition)
Rey Complex Figure Test (RCFT; Meyers & Meyers, 1995)	*Stimulus card* and *response paper* in an envelope at the participant's location; standard instructions provided with pointing omitted	• Copy time (in s) • Copy total raw score • Delay (3-min) total raw score
Wechsler Memory Scale – Fourth Edition (WMS-IV) Visual Reproduction (VR; Wechsler, 2009b)	*Stimulus books 1* and *2* and *response booklet* at the participant's location; research assistant changed pages in accordance with the examiner's instruction; standard instructions provided with pointing omitted	• Total raw score for: o VR I o VR II o VR Recognition
Executive function		
Letter Fluency (FAS; Strauss et al., 2006)	None required	• Number of correct responses
Stroop Test (Victoria Modification; Regard, 1981)	*Stimulus cards D, W,* and *C* in separate envelopes at the participant's location; standard instructions provided with pointing omitted	• Dots time (in s) • Words time (in s) • Colour Words time (in s) • Interference (i.e., Colour Words/Dots)
Trail Making Test (TMT; Reitan & Wolfson, 1985)	*TMT A* and *B response forms* in separate envelopes at the participant's location; standard instructions provided with pointing omitted; verbal error correction	• TMT A time (in s) • TMT B time (in s)
WAIS-IV Similarities (Wechsler, 2008)	None required	• Total raw score

SDMT = Symbol Digit Modalities Test; WAIS = Wechsler Adult Intelligence Scale; HVLT-R = Hopkins Verbal Learning Test – Revised.

are shown in Table 1. Raw scores, rather than standardised scores, were used for analyses as they have a greater range, allowing for a more nuanced comparison of performance. We used pairwise deletion of missing values for all analyses.

Comparing in-person and videoconference scores

A series of repeated-measures *t* tests were used to determine whether there was a significant difference between in-person and videoconference scores for each measure. Transformations were not conducted for non-normality as *t* tests are robust to violations of normality with sample sizes greater than 30 (Field, 2018). As we were looking for no difference between conditions (and therefore non-significant *t* tests), a less stringent alpha was more conservative in this instance (i.e., a Bonferroni adjustment was not applied).

In keeping with similar studies (e.g., Cullum et al., 2014), intraclass correlation coefficient (ICC) estimates were used to assess the reliability of repeated (i.e., in-person and videoconference) administrations. We used single occasion, absolute agreement, two-way random effects ICC estimates and their 95% confidence intervals (Koo & Li, 2016). Bujang and Baharum (2017) suggest that the minimum sample size requirement for estimating ICCs to assess the reliability of different measurement methods varies between 18 and 50, therefore, our sample size was sufficient for these analyses. Normality was assessed for all variables using converging evidence from visual inspection of histograms and standardised skewness and kurtosis values. For most variables where issues of non-normality were identified, winsorising outliers remedied or significantly reduced these issues (Field, 2018). Analysis of relevant ICC estimates pre- and post-winsorising indicated only minimal influence of outliers on the results for these measures. Further, in most instances, winsorising outliers deflated rather than inflated the relevant correlations. We did not transform distributions with remaining issues of mild non-normality, as this would have significantly confused the interpretation of results.

Bland–Altman plots were constructed to further evaluate the agreement between conditions (Bland & Altman, 1986). In a Bland–Altman plot, an individual's average score on a measure is plotted against their difference score on the measure (i.e., videoconference score minus in-person score; Bland & Altman, 1986). A Bland–Altman plot shows the bias (i.e., average difference) value and the 95% limits of agreement (i.e., limits within which 95% of difference scores will lie), both derived from average difference scores (Bland & Altman, 1986). In this study, for most measures, positive bias values represent superior performance in the videoconference condition, on average, while negative bias values represent superior performance in the in-person condition, on average. The opposite is true where higher numbers represent inferior performance on a measure (e.g., TMT). Winsorising outliers remedied non-normal difference distributions, where necessary (Field, 2018).

If converging evidence from the above analyses indicated differences in test performance, we conducted further analyses. In this instance, we used multivariable models to evaluate the influence of participant characteristics on this outcome.

Acceptability of administration methods

For items where participants had to rate their satisfaction, ease of understanding during sessions or comfort with the videoconference equipment (items 1–5), numeric values from 1 (indicating the least favourable response; e.g., *completely dissatisfied*) to 5 (indicating the most favourable response; e.g., *completely satisfied*) were applied to response options. Averages were calculated, with higher scores representing greater satisfaction, understanding, or comfort. All other question responses were summarised by endorsement frequencies and percentages.

RESULTS

Sample Characteristics

Figure 1 displays participant recruitment, progression through the study, and counterbalancing. The most frequently reported reason for declining to participate was time constraints (due to rehabilitation commitments, return to work, etc.). Table 2 presents the demographic and clinical characteristics of the 48 participants. Years of education were calculated using the norms of Heaton, Miller, Taylor, and Grant (2004). Participants were on average in their mid-60s, Australian born, and had had a stroke over 5 years previously. Most participants had experienced an ischaemic stroke (68.8%) and most had experienced a left hemisphere stroke (50%). Fewer participants had experienced a right hemisphere stroke (33.3%) or bilateral strokes (12.5%). Stroke location was unknown for two (4.2%) participants. Sessions were completed on average 15.8 ($SD = 9.7$) days apart.

Comparing In-person and Videoconference Scores

Due to participant fatigue and time constraints, not all participants completed all measures. Further, due to language impairment, one participant did not complete all measures. Table 3 presents the number of participants who completed each measure and means and standard deviations of scores in the videoconference and in-person conditions. Mean scores were similar across conditions for most tests. Most pairwise differences were not statistically significant (all *p*s > .05). However, pairwise comparisons indicated that, on average, participants remembered significantly fewer words across HVLT-R learning trails (i.e., Total Recall) in the videoconference condition than the in-person condition, $t(44) = 2.65$, $p = .011$, $d = 0.39$; there was a small effect size (Cohen, 1988). In addition, on average, Stroop Interference scores were superior in the videoconference condition with

Table 2. Demographic and clinical characteristics of the sample

	n (%)	*M (SD)*	Range
Age (years)	48 (100)	64.6 (10.1)	35–88
Sex (male)	26 (54.2)		
Education (years)	48 (100)	13.69 (3.3)	8–20
Country of birth	48 (100)		
Australia	33 (68.7)		
England	10 (20.8)		
Other	5 (10.4)		
MoCA	48 (100)	24.1 (3.4)	8.5–29.5
CPQ	44 (91.7)	22.0 (6.2)	6–30
HADS-A	45 (93.8)	5.9 (4.0)	0–16
Normal	27 (56.3)		
Mild	12 (25.0)		
Moderate	5 (10.4)		
Severe	1 (2.1)		
HADS-D	45 (93.8)	4.8 (3.8)	0–15
Normal	36 (75.0)		
Mild	6 (12.5)		
Moderate	2 (4.2)		
Severe	1 (2.1)		
Years since stroke	47 (97.9)	5.2 (4.0)	0.3–16.5
Stroke mechanism			
Ischemic	33 (68.8)		
Hemorrhagic	5 (10.4)		
Both	6 (12.5)		
Unknown	4 (8.3)		
Stroke hemisphere			
Left	24 (50.0)		
Right	16 (33.3)		
Bilateral	6 (12.5)		
Unknown	2 (4.2)		

MoCA = Montreal Cognitive Assessment; CPQ = Computer Proficiency Questionnaire; HADS-A = Hospital Anxiety and Depression Scale – Anxiety; HADS-D = Hospital Anxiety and Depression Scale – Depression.

a small effect size, $t(40) = 2.25$, $p = .030$, $d = 0.35$ (Cohen, 1988).

ICC estimates and their 95% confidence intervals are also shown in Table 3. ICC estimates ranged from .40 to .96 across measures. Tests with the highest ICC estimates included the TOPF, WAIS-IV Digit Span, and FAS. In contrast, ICC estimates for the HVLT-R Total Recall and Discrimination Index were lower (i.e., .47 and .40, respectively). While the ICC estimate for Stroop Interference scores was low (.49), ICC estimates for Stroop components, particularly Words and Colour Words, were higher (i.e., .71–.86).

Parameters for the Bland–Altman plots (i.e., bias values, upper and lower 95% limits of agreement) and their associated 95% confidence intervals are presented in Table 3. The Bland–Altman plots are available in the online Supplementary Material. Bias values were close to 0 for most measures. However, the bias value for the WMS-IV VR II indicated superior average performance in the videoconference condition. In addition, the bias value for HVLT-R Total Recall indicated superior average performance in the in-person condition.

Across measures, the 95% limits of agreement were relatively wide. In addition, most measures showed relative symmetry in points above and below 0, indicating participants did not perform better or worse in a particular condition. However, for the HVLT-R Total Recall, Delayed Recall, and Discrimination Index, more participants had negative difference values, indicating better performance in the in-person condition. In four Bland–Altman plots, there was an indication of unequal variance in difference values along the spectrum of average values (i.e., heteroscedasticity). For the HVLT-R Total Recall and Discrimination Index and the RCFT Copy score, those with lower average scores (indicating poorer performance) varied more across sessions than those with higher average scores. Similarly, those with higher average scores on TMT A (indicating poorer performance) varied more across sessions than those with lower average scores.

We ran further analyses to evaluate whether participant characteristics could explain the poorer performance on HVLT-R Total Recall in the videoconference condition (we assumed that results would be similar across HVLT-R scores, which showed a similar pattern of performance across conditions). In this multivariable regression analysis, participant characteristics included in the model (i.e., age, level of cognition [MoCA], computer proficiency [CPQ], and symptoms of anxiety [HADS-A] and depression [HADS-D]) did not significantly predict HVLT-R Total Recall difference scores, $F(5, 35) = 1.64$, $p = .175$, adj. $R^2 = .074$. Table 4 displays the regression coefficients. No predictors contributed to the model (all ps > .05).

Acceptability of Administration Methods

Forty-five participants completed the acceptability survey. Table 5 shows the average ratings for satisfaction, ease of understanding during each condition, and comfort with the videoconference equipment. For other items, endorsement frequencies and percentages are shown. Average satisfaction ratings were comparable across conditions and reflected that participants were, on average, satisfied with both conditions. The majority of respondents reported equal comfort in both conditions and reported no preference for a particular condition. Of those who preferred the in-person condition ($n = 19$), the majority suggested that this was because this condition facilitated a better interpersonal connection with the examiner ($n = 10$). This was also the most frequently endorsed advantage of the in-person session by all participants. Other reasons included that the in-person session had less scope for technical glitches ($n = 3$) and allowed the participant to better read the examiner's body language ($n = 3$). Of the three who preferred the videoconference condition, one participant started feeling more relaxed in this condition and another suggested that this type of interaction could eliminate future travel. In all, 24.4% of the sample reported the videoconference session as more interesting or fun. While the in-person session was the most preferred session, most participants reported being

Table 3. Means (*M*), standard deviations (*SD*) and ranges of in-person and videoconference scores, and ICCs and Bland–Altman parameters comparing in-person and videoconference scores

		In-person			Videoconference		ICC [95% CI]	Bland–Altman parameters		
			Range							
	n	*M (SD)*	Possible	Observed	*M (SD)*	Range		Bias [95% CI]	Lower LOA [95% CI]	Upper LOA [95% CI]
TOPF	47	50.51 (12.58)	0–70	16–67	50.89 (13.21)	21–68	.96 [.92, .97]	.38 [−.62, 1.39]	−7.17 [−8.93, −5.41]	7.94 [6.18, 9.70]
WAIS-IV Block Design	44	36.14 (10.73)	0–66	20–58	36.14 (10.30)	20–55	.74 [.57, .85]	.00 [−2.31, 2.31]	−14.97 [−18.98, −10.96]	14.97 [10.96, 18.98]
WAIS-IV Similarities	45	23.76 (4.88)	0–36	9–32	23.11 (5.43)	11–33	.87 [.78, .93]	−.64 [−1.41, .12]	−5.68 [−7.01, −4.35]	4.39 [3.05, 5.72]
WAIS-IV Digit Span	46	26.61 (6.99)	0–48	6–40	26.20 (7.22)	10–40	.88 [.79, .93]	−.41 [−1.47, .64]	−7.38 [−9.21, −5.56]	6.56 [4.73, 8.38]
DSF	46	10.09 (2.88)	0–16	4–16	9.93 (3.11)	4–16	.87 [.78, .93]	−.15 [−.60, .30]	−3.16 [−3.95, −2.37]	2.85 [2.06, 3.64]
DSB	46	8.89 (2.87)	0–16	0–16	8.43 (2.88)	4–15	.76 [.61, .86]	−.46 [−1.04, .12]	−4.30 [−5.31, −3.29]	3.39 [2.38, 4.40]
DSS	46	7.87 (2.38)	0–16	1–13	7.83 (2.48)	2–13	.70 [.52, .82]	−.04 [−.60, .51]	−3.74 [−4.71, −2.77]	3.65 [2.68, 4.62]
WMS-IV VR I	46	34.02 (6.34)	0–43	17–43	33.96 (5.75)	22–43	.73 [.57, .84]	−.06 [−1.38, 1.25]	−8.77 [−11.04, −6.49]	8.64 [6.37, 10.92]
WMS-IV VR II	46	22.61 (9.91)	0–43	0–41	23.98 (10.62)	0–42	.73 [.56, .84]	1.37 [−.86, 3.59]	−13.35 [−17.21, −9.50]	16.09 [12.24, 19.95]
WMS-IV VR Recognition	45	5.80 (1.27)	0–7	2–7	5.67 (1.36)	2–7	.71 [.52, .83]	−.13 [−.44, .17]	−2.12 [−2.64, −1.59]	1.85 [1.33, 2.38]
BNT	47	55.04 (4.81)	0–60	41–60	54.89 (4.70)	43–60	.86 [.76, .92]	−.17 [−.92, .58]	−5.18 [−6.48, −3.88]	4.84 [3.54, 6.14]
HVLT-R Total Recall	45	23.69 (5.55)	0–36	10–33	21.58 (5.13)	10–32	.47 [.21, .67]	−2.11 [−3.71, −.51]	−12.58 [−15.35, −9.80]	8.35 [5.58, 11.12]
HVLT-R Delayed Recall	44	7.70 (2.83)	0–12	0–12	7.04 (3.18)	0–12	.61 [.39, .76]	−.66 [−1.46, .14]	−5.82 [−7.20, −4.44]	4.50 [3.11, 5.88]
HVLT-R Discrimination Index	43	10.16 (1.51)	0–12	7–12	9.70 (2.17)	5–12	.40 [.12, .62]	−.46 [−1.09, .16]	−4.46 [−5.55, −3.37]	3.53 [2.44, 4.62]
Letter Fluency (FAS)	41	38.49 (17.06)	≥0	11–85	38.34 (16.86)	11–90	.89 [.80, .94]	−.15 [−2.67, 2.37]	−15.80 [−20.17, −11.44]	15.51 [11.15, 19.88]
RCFT Copy time (s)	39	165.74 (53.95)	≥1	77–335	169.59 (56.86)	88–270	.61 [.37, .78]	6.54 [−6.65, 19.73]	−73.38 [−96.22, −50.53]	86.45 [63.61, 109.30]
RCFT Copy score	42	30.06 (3.73)	0–36	19.5–35	30.40 (3.40)	20.5 –35	.74 [.57, .85]	.14 [−.79, 1.07]	−5.71 [−7.33, −4.10]	6.00 [4.39, 7.61]
RCFT Delay score	41	17.02 (7.23)	0–36	2.50–31	16.63 (7.12)	2.50 –31	.80 [.65, .89]	−.39 [−1.84, 1.06]	−9.42 [−11.94, −6.91]	8.65 [6.13, 11.17]
Semantic Fluency (Animals)	41	18.83 (4.99)	0 +	11–32	18.76 (5.75)	10–34	.68 [.48, .82]	−.22 [−1.45, 1.00]	−7.81 [−9.92, −5.69]	7.36 [5.24, 9.47]
SDMT	47	45.02 (10.42)	0–110	24–68	45.17 (10.07)	21 –71	.82 [.70, .90]	.15 [−1.65, 1.95]	−11.88 [−14.99, −8.76]	12.18 [9.05, 15.29]
Stroop Test Interference	41	2.16 (.59)	–	1.39–3.27	1.96 (.49)	1.16–2.90	.49 [.22, .69]	−.20 [−.37, −.03]	−1.24 [−1.53, −.95]	.84 [.55, 1.13]
Dots (s)	41	16.00 (4.14)	≥1	9–26	17.98 (5.97)	10–34	.71 [.45, .84]			
Words (s)	41	21.41 (7.79)	≥1	11–44	21.10 (6.36)	13–41	.88 [.78, 93]			
Colour Words (s)	41	35.90 (14.00)	≥1	18–70	35.07 (13.59)	17–72	.86 [.76, .92]			
TMT A (s)	47	39.02 (14.14)	≥1	19–72	41.51 (13.03)	21–67	.69 [.51, .82]	2.85 [−.50, 6.21]	−19.60 [−25.42, −13.78]	25.31 [19.49, 31.12]
TMT B (s)	47	105.04 (44.91)	≥1	41–195	108.21 (54.66)	37–284	.85 [.74, .91]	2.15 [−4.95, 9.25]	−45.32 [−57.62, −33.02]	49.62 [37.32, 61.92]

ICC = intraclass correlation coefficient; CI = confidence interval; LOA = limit of agreement; TOPF = Test of Premorbid Function; WAIS-IV = Weschler Adult Intelligence Scale – Fourth Edition; DSF = Digit Span Forward; DSB = Digit Span Backward; DSS = Digit Span Sequencing; WMS-IV VR = Weschler Memory Scale – Fourth Edition Visual Reproduction; BNT = Boston Naming Test; HVLT-R = Hopkins Verbal Learning Test – Revised; RCFT = Rey Complex Figure Test; SDMT = Symbol Digit Modalities Test; TMT = Trail Making Test.

Table 4. Unstandardised (*B*) and Standardised (*β*) regression coefficients, and semi-partial correlations (*Sr²*) of predictors in the model predicting HVLT-R Total Recall difference scores

	Unstandardised coefficients				
	B	*SE*$_B$	95% CI	*β*	*Sr²*
Intercept	13.26	6.06	[.96, 25.57]		
Age	−.83	.45	[−.17, .01]	−.32	−.28
MoCA	−.24	.18	[−.61, .12]	−.24	−.21
CPQ total	−.13	.07	[−.28, .02]	−.28	−.26
HADS-A	.13	.12	[−.12, .38]	.19	.16
HADS-D	−.15	.13	[−.41, .12]	−.20	−.17

HVLT-R = Hopkins Verbal Learning Test – Revised; MoCA = Montreal Cognitive Assessment; CPQ = Computer Proficiency Questionnaire; HADS-A = Hospital Anxiety and Depression Scale – Anxiety; HADS-D = Hospital Anxiety and Depression Scale – Depression.

unwilling to wait for more than 3 months or travel long distances for this type of assessment.

DISCUSSION

We report the first comparison of performance of a comprehensive neuropsychological battery administered in-person and *via* videoconference, and the first evaluation of the acceptability of videoconference-based neuropsychological assessment among stroke survivors. To our knowledge, we are the first to compare results across in-person and videoconference-based administrations for several neuropsychological measures including the TOPF, RCFT, Stroop Test, WAIS-IV Block Design, WAIS-IV Similarities, and WMS-IV VR. For most measures, converging evidence indicated that participants did not perform systematically better or worse in a particular condition. Therefore, our study provides preliminary evidence that test results across in-person and videoconference-based administrations are comparable for community-based survivors of stroke and could potentially be used interchangeably in clinical practice for this group. Inclusion of the Bland–Altman limits of agreement gives clinicians a resource to evaluate relative confidence in the comparability of results for each measure. In contrast, converging evidence indicated that participants performed more poorly on the HVLT-R in the videoconference condition than the in-person condition. In addition, Stroop Test Interference scores were superior in the videoconference condition than the in-person condition. This indicates that these measures should potentially be avoided, or appropriate considerations should be made (e.g., conservative clinical decision-making), when using these measures *via* videoconference in clinical practice. We also found that participants were broadly accepting of videoconference-based neuropsychological assessment and were prepared to avoid travel and delays in access to a neuropsychologist as a trade-off for in-person assessments.

Examining previous research, the authors of a recent meta-analysis evaluating agreement of in-person and videoconference neuropsychological scores across previous studies using non-stroke samples (e.g., healthy participants, those with dementia) have demonstrated that videoconference scores were one-tenth of a standard deviation below in-person scores for non-timed tests that allow for repetition, which they defined as *nonsynchronous dependent tests* (Brearly et al., 2017). The current results are broadly consistent with these findings. For example, the BNT demonstrated an ICC estimate of .86 and the RCFT Delay score had an ICC estimate of .80. In this meta-analysis, it was also shown that in-person and videoconference scores were equivalent for *synchronous dependent* tasks, which are verbally mediated, timed tests that proscribe repetition (e.g., list learning tasks; Brearly et al., 2017). Interestingly, while we did demonstrate equivalent scores for several of these synchronous dependent tasks, for example, WAIS-Digit Span, verbal fluency tasks, and SDMT, the HVLT-R was not equivalent in our study. This may, in part, reflect the test–retest reliability of this measure broadly, which is shown to be sub-optimal, particularly for Delayed Recall (*r* = .66) and the Discrimination Index (*r* = .40; Benedict, Schretlen, Groninger, & Brandt, 1998), in comparison with other measures in this study. Indeed, it should be noted that most of the ICC estimates in our study did not meet the acceptable limit of ≥.90 for a psychological research context, as defined by Nunnally and Bernstein (1994). However, as discussed above, the participants did not perform systematically better/worse in a particular condition. Instead, current findings likely reflect the broader reliability of the individual measures we were evaluating. Indeed, for other measures, ICC estimates in our study seemed to match established reliability coefficients. For example, TMT A had a lower ICC estimate than TMT B, which is consistent with test–retest reliability estimates established for this measure in the literature (Strauss et al., 2006). Further, the WAIS-IV subtests evaluated had similar ICC estimates to established test–retest reliability coefficients presented in the WAIS-IV test manual (Wechsler, 2008). However, participants in our sample did, on average, perform worse on the HVLT-R in the videoconference condition specifically, which would not be expected solely on the basis of poor test–retest reliability. This difference was not explained by participant characteristics such as age, cognitive impairment, computer proficiency, or depression or anxiety symptoms. This might have reflected the dependence of this test in particular on the highly synchronous transfer of both visual and verbal cues, which may have meant that it was harder to hear words in the videoconference condition or that there was a higher chance of mishearing words in this condition. It is also possible that due to this fact, participants were particularly anxious about, or preoccupied with, the videoconference scenario for this test, which could have affected their performance. It should be an aim of future research to replicate this finding, particularly in those with a more diverse range of cognitive abilities. Future research could also evaluate whether the use of a pre-recorded word list (rather than examiner reading), or verbal memory tasks that also present the written word, results in a similar trend. In addition,

Table 5. Means (*M*), standard deviations (*SD*), and endorsement frequencies for items on the acceptability measure

Question	*M* (*SD*)	Range	*n* (%)
Satisfaction rating[a]			
In-person condition	4.8 (0.7)	2–5	
Videoconference condition	4.7 (0.8)	1–5	
Understanding of task instructions[a]			
In-person condition	4.8 (0.4)	4–5	
Videoconference condition	4.6 (0.8)	2–5	
Comfort with videoconference equipment[a]	4.6 (0.8)	1–5	
Willingness to recommend videoconference services			
Yes			40 (88.9)
No			3 (6.6)
Preferred condition			
In-person			19 (42.2)
No preference			23 (51.1)
Videoconference			3 (6.6)
Willingness to travel for in-person consultation			
Less than 1 h			15 (33.3)
1–3 h			13 (28.9)
3–6 h			2 (4.4)
As far as necessary			2 (4.4)
Preference for videoconference testing			12 (26.7)
Willingness to wait for in-person consultation			
Less than 1 month			13 (28.9)
1–3 months			9 (20.0)
3–6 months			1 (2.2)
As long as it takes			2 (4.4)
Preference for videoconference testing			17 (37.8)
Prefer videoconference over less qualified professional in-person			
Yes			29 (64.4)
No			12 (26.7)
Comfort ranking			
More comfortable in in-person session			10 (22.2)
Equal comfort			33 (73.3)
More comfortable in videoconference session			2 (4.4)
Advantages of the in-person condition[b]			
Easier to establish personal connection with examiner			30 (66.7)
Easier to communicate with the examiner			19 (42.2)
Easier to understand use of test materials			14 (31.1)
Videoconference equipment had poor quality sound			4 (8.9)
Hard to hear the examiner in the videoconference condition			6 (13.3)
Videoconference equipment had poor visual quality			2 (4.4)
Hard to see the examiner in the videoconference condition			1 (2.2)
Harder to comprehend instructions in videoconference condition			5 (11.1)
Other (e.g., less scope for technical glitches)			4 (8.9)
Advantages of videoconference condition[b]			
Easier to establish personal connection with examiner			6 (13.3)
Easier to communicate with the examiner			3 (6.6)
Less anxious/nervous without examiner in the room			4 (8.9)
Easier to concentrate without examiner in the room			4 (8.9)
More interesting/fun			11 (24.4)
Felt more in control			4 (8.9)
Other (e.g., would reduce travel, new experience)			8 (17.8)

[a] Possible ratings range from 1 (*least favourable*) to 5 (*most favourable*).
[b] Participants selected all options that applied.

further research could evaluate whether other verbal learning tasks (with better reliability and semantic content; e.g., story memory tasks) result in more comparable results across conditions.

Difference was also demonstrated between conditions for the Stroop Test Interference score, with the videoconference condition having superior (i.e., lower) scores. However, it seems that this difference was actually driven by marginally slower performance on the Dots trial in the videoconference condition. Interestingly, this finding was isolated to this trial, with both the Words and Colour Words trials having largely similar average results across conditions. While this may be a spurious finding, it may also be related to the fact that Dots is the first trial to be administered. That is, it is possible that participants were particularly concerned about the examiner's capacity to hear them in this trial. Further research is needed to replicate this finding. Another finding was that the RCFT Copy and TMT A demonstrated more variable difference scores in those who were performing more poorly on these measures. However, this may reflect the variable psychometric properties of neuropsychological measures for people with different levels of cognitive function.

Our results regarding the acceptability of videoconference-based neuropsychological assessment were broadly consistent with previous research in other samples. Parikh et al. (2013) had 40 participants who were either healthy or had diagnoses of AD or MCI complete an acceptability survey following neuropsychological assessment both in-person and *via* videoconference. Their results reflect 98% satisfaction with videoconference-based neuropsychological assessment, which was consistent with the high average satisfaction rating (4.7 out of 5) reported in our sample. In addition, in their sample, 60% of participants had no preference for a particular session, 30% preferred in-person assessment, and 10% preferred videoconference-based assessment (Parikh et al., 2013). These findings are broadly in keeping with our findings, with a slightly higher percentage of our sample (42.2%) reporting a preference for in-person assessment. Beyond Parikh et al.'s (2013) findings, similar rates of preference across conditions have been reported in other studies (i.e., Hildebrand et al., 2004; Jacobsen et al., 2003; Stain et al., 2011). Whilst interpersonal connection was the main driver of the preference for in-person assessments in our sample, it did not seem to outweigh the burden of travel or wait times. The vast majority of participants in our sample indicated that they would prefer videoconference-based consultations if it avoided travel of more than 3 hours or a wait time of more than 3 months. One limitation of our evaluation of acceptability is that the evaluation occurred after the second session. Whilst this was necessary to facilitate comparison of the conditions, this also may have been difficult for some participants with memory difficulties given the 2- week interval between sessions. Perhaps, a more ecologically valid measure of acceptability would be gained by assessing the acceptability of each condition independently, directly after session completion. Further, as noted in Figure 1, a number of people declined to participate in this study. Whilst not

explicitly stated by these individuals, it is possible that the use of videoconferencing in this study deterred them from participating. As such, this may have served to increase evaluations of acceptability in this study.

This study has several strengths. First, we included neuropsychological measures that facilitated the assessment of all classically assessed domains of cognitive function. Some previous studies in non-stroke populations have limited their batteries to verbal tasks, which are particularly suitable to administration *via* videoconference. This, however, limits the assessment of some neuropsychological domains, particularly visuospatial function and nonverbal problem-solving, and therefore the comprehensiveness of the neuropsychological evaluation being validated. Second, the condition order was counterbalanced in the current study, as well as the HVLT-R form version, which reduced the potential influence of practice effects. Third, we used low-cost, easily accessible hardware and low-cost, easily accessible, and secure videoconference software. These features were considered important to maximise the likelihood of clinical translation. Most features of this setup would be readily available in health services or would require minimal funding. In addition, most healthcare providers, and possibly patients, would be familiar with this hardware and software.

This study also has a number of limitations. First, the study had the potential for bias due to the use of a single examiner. However, it should also be noted that the use of separate examiners could introduce the potential for bias due to inter-examiner differences. Further research should aim to counterbalance different examiners alongside the condition of participants. Second, we did not include a condition in which participants completed the neuropsychological measures in-person in both sessions. This would have provided a clearer comparison for the reliability and agreement statistics presented here. Third, whilst the inclusion of a comprehensive battery of neuropsychological measures was a clear strength of this study, we did not include a measure of hemispatial neglect (despite the prevalence of this syndrome following stroke) nor an established test battery that allows for profile analysis. These should be included in future research. The use of a comprehensive battery also necessitated the use of a research assistant to facilitate administration of some neuropsychological measures requiring a higher level of examiner control over stimulus materials (e.g., WAIS-IV Block Design). This administration modification was necessary to ensure the protection of copyright for these materials. However, this also may present additional barriers in terms of clinical translation. It should be the objective of future work to consider alternatives to test administration that do not require an assistant to be present, perhaps in consultation with the publishers of these tests. Finally, the participant sample was on average 5 years post-stroke, was relatively computer proficient, and mostly Australian born. In addition, because we used a community sample, we were unable to obtain a measure of stroke severity. However, MoCA results did indicate that our participants were, on average, only mildly cognitively impaired. The findings presented herein

may not be generalised beyond the relatively mildly impaired population of stroke survivors evaluated in the current study. Future research should aim to replicate the above findings in acute and subacute samples, in those who have more severe cognitive impairments and are more representative of the diverse stroke population, and in those that are less computer proficient. Further, future research should also evaluate the suitability of videoconference-based assessment for those with more severe cognitive impairments and/or behavioural disturbance who may present additional challenges when assessed in a videoconference context. Of course, the administration of neuropsychological tests is only one, albeit central element of neuropsychological practice. As such, future research should also be designed to evaluate other aspects of neuropsychological practice, including clinical interviewing, assessing client presentation, intervention, and secondary consultation to other disciplines, *via* videoconference.

This study provides preliminary evidence to support videoconference-based administration of a number of common neuropsychological tasks in community-based survivors of stroke. Whilst further research in this area is warranted, particularly with regard to the HVLT-R, videoconference-based neuropsychological assessment stands to have substantial benefits for improving access to neuropsychological assessment and treatment for stroke. Indeed, the importance and value of telehealth have been highlighted in recent times, where restrictions put in place by the COVID-19 pandemic have limited the capacity of neuropsychologists around the world to engage in in-person assessment practices. Further, the conduct of neuropsychological assessments *via* videoconference may attenuate or eliminate some of the regional disparities in the availability of neuropsychological services.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S1355617720001174.

REFERENCES

Australian Institute of Health and Welfare. (2018). *Australia's Health 2018. (Australia's Health Series No. 16. AUS 221).* Canberra, ACT: AIHW.

Bartlett, J. & Wetzel, R. (2010). Video conferencing quality vs. bandwidth tradeoffs. Retrieved from http://searchunifiedcommunications.techtarget.com/tip/Video-conferencing-quality-vs-bandwidth-tradeoffs

Benedict, R.H.B., Schretlen, D., Groninger, L., & Brandt, J. (1998). Hopkins Verbal Learning Test – Revised: Normative data and analysis of inter-form and test-retest reliability. *The Clinical Neuropsychologist*, 12(1), 43–55. doi: 10.1076/clin.12.1.43.1726

Bjelland, I., Dahl, A.A., Haug, T.T., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale: An updated literature review. *Journal of Psychosomatic Research*, 52(2), 69–77. doi: 10.1016/s0022-3999(01)00296-3

Bland, J.M. & Altman, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 8(1), 307–310. doi: 10.1016/S0140-6736(86)90837-8

Boot, W.R., Charness, N., Czaja, S.J., Sharit, J., Rogers, W.A., Fisk, A.D., . . . Nair, S. (2015). Computer Proficiency Questionnaire: Assessing low and high computer proficient seniors. *The Gerontologist*, 55(3), 404–411. doi: 10.1093/geront/gnt117

Brandt, J. & Benedict, R.H.B. (2001). *Hopkins Verbal Learning Test – Revised*. Odessa, TX: Psychological Assessment Resources.

Brearly, T.W., Shura, R.D., Martindale, S.L., Lazowski, R.A., Luxton, D.D., Shenal, B.V., & Rowland, J.A. (2017). Neuropsychological test administration by videoconference: A systematic review and meta-analysis. *Neuropsychological Review*, 27(2), 174–186. doi: 10.1007/s11065-017-9349-1

Bujang, M.A. & Baharum, N. (2017). A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review. *Archives of Orofacial Sciences*, 12(1), 1–11. Retrieved from http://www.dental.usm.my/aos/docs/Vol_12/aos-article-0246.pdf

Burton, L. & Tyson, S.F. (2015). Screening for cognitive impairment after stroke: A systematic review of psychometric properties and clinical utility. *Journal of Rehabilitation Medicine*, 47(3), 193–203. doi: 10.2340/16501977-1930

Burton, R.L. & O'Connell, M.E. (2018). Telehealth rehabilitation for cognitive impairment: Randomized controlled feasibility trial. *JMIR Research Protocols*, 7(2), e43. doi: 10.2196/resprot.9420

Chapman, J.E., Cadilhac, D.A., Gardner, B., Ponsford, J., Bhalla, R., & Stolwyk, R.J. (2019). Comparing face-to-face and videoconference completion of the Montreal Cognitive Assessment (MoCA) in community-based survivors of stroke. *Journal of Telemedicine and Telecare*. Advance online publication. doi: 10.1177/1357633X19890788

Chapman, J.E., Ponsford, J., Bagot, K.L., Cadilhac, D.A., Gardner, B., & Stolwyk, R.J. (2020). The use of videoconferencing in clinical neuropsychological practice: A mixed methods evaluation of neuropsychologists' experiences and views. *Australian Psychologist*. Advance online publication. doi: 10.1111/ap.12471

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cullum, C.M. & Grosch, M.C. (2013). Special considerations in conducting neuropsychology assessment over videoconferencing. In K. Myers & C.L. Turvey (Eds.), *Telemental health clinical, technical, and administrative foundations for evidenced based practice* (pp. 275–293). doi: 10.1016/B978-0-12-416048-4.00014-2n

Cullum, C.M., Hynan, L.S., Grosch, M.C., Parikh, M., & Weiner, M.F. (2014). Teleneuropsychology: Evidence of video teleconference-based neuropsychological assessment. *Journal of the International Neuropsychological Society*, *20*(10), 1–6. doi: 10.1017/S1355617714000873

Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (5th ed.). London, England: SAGE Publications Ltd.

Hachinski, V., Iadecola, C., Peterson, R.C., Breteler, M.M., Nyenhuis, D.L., Black, S.E., … Leblanc, G.G. (2006). National Institute of Neurological Disorders and Stroke-Canadian Stroke Network vascular cognitive impairment harmonization standards. *Stroke*, *37*(9), 2220–2241. doi: 10.1161/01.STR.0000237236.88823.47

Heaton, R.K., Miller, S.W., Taylor, M.J., & Grant, M.J. (2004). *Revised Comprehensive Norms for An Expanded Halstead-Reitan Battery: Demographically Adjusted Neuropsychological Norms for African American and Caucasian Adults.* Lutz, FL: Psychological Assessment Resources.

Hildebrand, R., Chow, H., Williams, C., Nelson, M., & Wass, P. (2004). Feasability of neuropsychological testing of older adults via videoconference: Implications for assessing the capacity for independent living. *Journal of Telemedicine and Telecare*, *10*(3), 130–134. doi: 10.1258/13576330432070751

Hodge, M.A., Sutherland, R., Jeng, K., Bale, G., Batta, P., Cambridge, A., … Silove, N. (2019). Agreement between telehealth and face-to-face assessment of intellectual ability in children with specific learning disorder. *Journal of Telemedicine and Telecare*, *25*(7), 431–437. doi: 10.1177/1357633X18776095.

Jacobsen, S.E., Sprenger, T., Andersson, S., & Krogstad, J.M. (2003). Neuropsychological assessment and telemedicine: A preliminary study examining the reliability of neuropsychological services performed via telecommunication. *Journal of the International Neuropsychological Society*, *9*(3), 472–478. doi: 10.1017/S1355617703930128

Janzen, L.A. & Guger, S. (2016). Clinical neuropsychology practice and training in Canada. *The Clinical Neuropsychologist*, *30*(8), 1193–1206. doi: 10.1080/13854046.2016.1175668

Kaplan, E.F., Goodglass, H., & Weintraub, S. (2001). *The Boston Naming Test* (2nd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Kirkwood, K., Peck, D.F., & Bennie, L. (2000). The consistency of neuropsychological assessments performed via telecommunication and face to face. *Journal of Telemedicine and Telecare*, *6*(3), 147–151. doi: 10.1258/13576330001935239

Koo, T.K. & Li, M.Y. (2016). A guideline for selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. doi: 10.1016/j.jcm.2016.02.012

Lawson, D.W., Stolwyk, R.J., Ponsford, J.L., McKenzie, D.P., Downing, M.G., & Wong, D. (2020). Telehealth delivery of memory rehabilitation following stroke. *Journal of the International Neuropsychological Society*, *26*(1), 58–71. doi: 10.1017/S1355617719000651.

Lesniak, M., Bak, T., Czepiel, W., Seniow, J., & Czlonkowska, A. (2008). Frequency and prognostic value of cognitive disorders in stroke patients. *Dementia and Geriatric Cognitive Disorders*, *26*(4), 356–363. doi: 10.1159/000162262

Marra, D.E., Hamlet, K.M., Bauer, R.M., & Bowers, D. (2020). Validity of teleneuropsychology for older adults in response to COVID-19: A systematic and critical review. *The Clinical Neuropsychologist*. Advance online publication. doi: 10.1080/13854046.2020.1769192

Martin-Khan, M., Flicker, L., Wootton, R., Loh, P.K., Edwards, H., Varghese, P., … Gray, L.C. (2012). The diagnostic accuracy of telegeriatrics for the diagnosis of dementia via video conferencing. *Journal of the American Medical Directors Association*, *13*(5), 19–24. doi: 10.1016/j.jamda.2012.03.004.

Martin-Khan, M., Varghese, P., Wootton, R., & Gray, L. (2007). Successes and failures in assessing cognitive function in older adults using video consultation. *Journal of Telemedicine and Telecare*, *13*(S3), 60–62. doi: 10.1258/135763307783247211

Meyers, J. & Meyers, K. (1995). *Rey Complex Figure Test and Recognition Trial: Professional Manual.* Lutz, FL: Psychological Assessment Resources.

Nasreddine, Z.S., Phillips, N.A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. doi: 10.1111/j.1532-5415.2005.53221.x

Nunnally, J.C. & Bernstein, I.H. (1994). *Psychometric Theory* (3rd ed.). New York, NY: McGraw-Hill Inc.

Nys, G.M.S., van Zandvoort, M.J.E., van der Worp, H.B., de Haan, E.H.F., de Kort, P.L.M., Jansen, B.P.W., & Kappelle, L.J. (2006). Early cognitive impairment predicts long-term depressive symptoms and quality of life after stroke. *Journal of the Neurological Sciences*, *247*(2), 149–156. doi: 10.1016/j.jns.2006.04.005

Parikh, M., Grosch, M.C., Graham, L.L., Hynan, L.S., Weiner, M.F., Shore, J.H., & Cullum, C.M. (2013). Consumer acceptability of brief videoconference-based neuropsychological assessment in older individuals with and without cognitive impairment. *The Clinical Neuropsychologist*, *27*(5), 808–817. doi: 10.1080/13854046.2013.791723

Pendlebury, S.T., Mariz, J., Bull, L., Mehta, Z., & Rothwell, P.M. (2012). MoCA, ACE-R, and MMSE versus the National Institute of Neurological Disorders and Stroke-Canadian Stroke Network Vascular Cognitive Impairment Harmonization Standards neuropsychological battery after TIA and stroke. *Stroke*, *43*(2), 464–469. doi: 10.1161/STROKEAHA.111.633586

Pendlebury, S.T., Mariz, J., Bull, L., Mehta, Z., & Rothwell, P.M. (2013). Impact of different operational definitions on mild cognitive impairment rate and MMSE and MoCA performance in transient ischaemic attack and stroke. *Cerebrovascular Disorders* *3*(5–6), 355–362. doi: 10.1159/000355496

Ponsford, J. (2016). The practice of clinical neuropsychology in Australia. *The Clinical Neuropsychologist*, *30*(8), 1179–1192. doi: 10.1080/13854046.2016.1195015

Psychology Board of Australia. (2018). *Psychology Board of Australia registrant data.* Retrieved from http://www.psychologyboard.gov.au/documents/default.aspx?record=WD18%2F25212&dbid=AP&chksum=z3QxEh3hWZCH5cOBGPQVAQ%3D%3D

Rebchuk, A.D., Deptuck, H.M., O'Neill, Z.R., Fawcett, D.S., Silverberg, N.D., & Field, T.S. (2019). Validation of a novel telehealth administration protocol for the NIH toolbox-cognitive battery. *Telemedicine and e-Health*, *25*(3), 237–242. doi: 10.1089/tmj.2018.0023

Regard, M. (1981). *Cognitive rigidity and flexibility: A neuropsychological study.* Unpublished Ph.D. Dissertation: University of Victoria.

Reitan, R.M. & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery: Therapy and Clinical Interpretation.* Tucson, AZ: Neuropsychological Press.

Saxena, S.K., Ng, T.P., Koh, G., Yong, D., & Fong, N.P. (2007). Is improvement in impaired cognition and depressive symptoms in post-stroke patients associated with recovery in activities of daily living? *Acta Neurologica Scandinavica*, *115*(5), 339–346. doi: 10.1111/j.1600-0404.2006.00751.x

Schopp, L.H., Johnstone, B.R., & Merveille, O.C. (2000). Multidimensional telecare stratergies for rural residents with brain injury. *Journal of Telemedicine and Telecare*, *6*(S1), 146–149. doi: 10.1258/1357633001934474

Skilbeck, C.E., Wade, D.T., Hewer, R.L., & Wood, V.A. (1983). Recovery after stroke. *Journal of Neurology, Neurosurgery, and Psychiatry*, *46*(1), 5–8. doi: 10.1136/jnnp.46.1.5

Smith, A. (1973). *Symbol Digit Modalities Test*. Los Angeles, CA: Western Psychological Services.

Stain, H.J., Payne, K., Thienel, R., Michie, P., Carr, V., & Kelly, B. (2011). The feasability of videoconferencing for neuropsychological assessments in rural youth experiencing early psychosis. *Journal of Telemedicine and Telecare*, *17*(6), 328–331. doi: 10.1258/jtt.2011.101015

Strauss, E., Sherman, E.M.S., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms and Commentary* (3rd ed.). New York, NY: Oxford University Press Inc.

Stroke Foundation. (2017). (Australian) Clinical Guidelines for Stroke Management 2017. Retrieved from https://strokefoundation.org.au/what-we-do/treatment-programs/clinical-guidelines

Stroke Foundation. (2018). National Stroke Audit Rehabilitation Services Report 2018. Retrieved from https://informme.org.au/stroke-data/Rehabilitation-audits

Stroke Foundation. (2019). National Stroke Audit Acute Services Report 2019. Retrieved from https://informme.org.au/stroke-data/Acute-audits

Sweet, J.J., Benson, L.M., Nelson, N.W., & Moberg, P.J. (2015). The American Academy of Clinical Neuropsychology, National Academy of Neuropsychology, and Society for Clinical Neuropsychology (APA Division 40) 2015 TCN Professional Practice and Salary Survey: Professional practices, beliefs, and incomes

of U.S. neuropsychologists. *The Clinical Neuropsychologist*, *29*(8), 1069–1162. doi: 10.1080/13854046.2016.1140228

Temple, V., Drummond, C., Valiquette, S., & Jozsvai, E. (2010). A comparison of intellectual assessments over video conferencing and in-person for individuals with ID: Preliminary data. *Journal of Intellectual Disability Research*, *54*(6), 573–577. doi: 10.1111/j.1365-2788.2010.01282.x.

Wadsworth, H.E., Dhima, K., Womack, K.B., Hart, J., Weiner, M.F., Hynan, L.S., & Cullum, C.M. (2018). Validity of teleneuropsychological assessment in older patients with cognitive disorders. *Archives of Clinical Neuropsychology*, *33*(8), 1040–1045. doi: 10.1093/arclin.acx140

Wadsworth, H.E., Galusha-Glasscock, J.M., Womack, K.B., Quiceno, M., Weiner, M.F., Hynan, L.S., … Cullum, C.M. (2016). Remote neuropsychological assessment in rural American Indians with and without cognitive impairment. *Archives of Clinical Neuropsychology*, *31*(5), 420–425. doi: 10.1093/arclin/acw030

Wagle, J., Farner, L., Flekkoy, K., Bruun Wyller, T., Sandvik, L., Fure, B., … Engedal, K. (2011). Early post-stroke cognition in stroke rehabilitation patients predicts functional outcome at 13 months. *Dementia and Geriatric Cognitive Disorders*, *31*(5), 379–387. doi: 10.1159/000328970

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale – Third Edition, Australian and New Zealand Language Adapted Edition.* Sydney, Australia: Pearson Clinical and Talent Assessment.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale – Fourth Edition, Australian and New Zealand Language Adapted Edition.* Sydney, Australia: Pearson Clinical and Talent Assessment.

Wechsler, D. (2009a). *Advanced Clinical Solutions for the WAIS-IV and WMS-IV.* San Antonio, TX: Pearson Clinical and Talent Assessment.

Wechsler, D. (2009b). *Wechsler Memory Scale – Fourth Edition, Australian and New Zealand Language Adapted Edition.* Sydney, Australia: Pearson Clinical and Talent Assessment.

Zigmond, A.S. & Snaith, R.P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361–370. doi: 10.1111/j.1600-0447.1983.tb09716.x

Zoom [Computer Software]. San Jose, CA: Zoom Video Communications Inc.