

Research Article

Telepractice of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): Validation and Practical Considerations

Carla Tortora , **Dalila Maglio** , **Irene Ceccato** , **Pasquale La Malva** ,
Adolfo Di Crosta , **Giulia Prete** , **Nicola Mammarella** , **Alberto Di Domenico** ,
and Rocco Palumbo 

Department of Psychology, University "G. d'Annunzio" of Chieti-Pescara, Chieti, Italy

Correspondence should be addressed to Carla Tortora; carla.tortora@unich.it

Received 9 January 2025; Revised 9 June 2025; Accepted 14 July 2025

Academic Editor: Daiana Colledani

Copyright © 2025 Carla Tortora et al. Human Behavior and Emerging Technologies published by John Wiley & Sons Ltd. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Telepractice in neuropsychology has become increasingly prevalent in recent years due to its ability to provide accessible and convenient care to patients regardless of their location. However, the validation of many neuropsychological tools for distance assessments remains limited, and there is a particular lack of remotely administered assessment tests with alternate forms, which are crucial for monitoring symptoms and performance in clinical contexts and for minimizing practice effects in research practice. Consequently, the present study was aimed at evaluating the consistency of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) scores across videoconference and face-to-face administrations and to examine whether the scores obtained via videoconference support interpretations similar to those obtained via face-to-face administration. A total of 185 participants aged between 20 and 79 years ($M = 46.24$, $SD = 19.63$) underwent RBANS testing twice: once in person using the standard pen-and-paper modality and once remotely via videoconference, using Alternate Forms A and B to mitigate the learning effects. Results from the linear mixed models revealed no significant differences between remote and face-to-face administrations based on the modality of administration ($p > 0.05$). Bayes factors supported the null hypothesis, suggesting that RBANS performance is consistent across the two modalities of administration. However, discrepancies were observed in certain subtests between alternate forms of the RBANS, highlighting the need for standardization. In conclusion, findings suggested that the same norms that are used to interpret the RBANS scores obtained via face-to-face administration may be employed when administered remotely through videoconferencing. Accordingly, the study provides valuable insights into the feasibility of remote neuropsychological assessment and underscores the potential utility of videoconference technology in clinical and research settings.

Keywords: aging; cognitive; cognitive testing; dementias; eHealth; neuropsychological test; telehealth

1. Introduction

Neuropsychological assessment encompasses a comprehensive evaluation of cognitive, behavioral, and emotional functioning that usually combines clinical interviews with standardized tests administered in person by trained professionals. While fundamental to understanding cognitive health, the conventional in-person administration of these

tests presents significant accessibility challenges, particularly for individuals residing in remote areas or facing limited mobility [1]. Given the importance of early screening and diagnosis, and to overcome these accessibility issues, there is a growing need for innovative approaches to ensure equitable access to comprehensive neuropsychological evaluations. Telemedicine, particularly teleneuropsychology, has emerged as a promising solution to overcome these barriers

[2]. Utilizing videoconferencing (VCF) and digital platforms, teleneuropsychology provides a means to remotely administer standardized tests that enable clinicians to assess cognitive functioning while resembling in-person evaluations.

VCF allows for real-time interactions between assessors and patients, potentially maintaining the integrity of the assessment process [3]. However, it introduces unique considerations. Technological limitations, the standardization of remote test administration protocols, and concerns regarding the validity and reliability of remote testing compared to in-person evaluations necessitate careful exploration [4]. Additionally, ensuring patient privacy, data security, and establishing standardized procedures for remote assessment administration are critical aspects requiring meticulous attention [5]. Moreover, a limitation of teleneuropsychology arises from potential usability challenges for older adults, as limited familiarity or access to the necessary technological devices may impede their participation in remote cognitive assessments and interventions [6]. Nonetheless, initial research indicates a rising percentage of older adults adapting to and using new technologies, suggesting a potential shift in their technological proficiency [7]. Preliminary results from early studies also show promising indications of the applicability and feasibility of teleneuropsychology for older adults with varying levels of cognitive impairment, emphasizing its potential usefulness within this population [3, 4, 8].

The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) [9] is a widely used cognitive assessment tool in both clinical and research settings. It was developed and refined to provide a brief but comprehensive evaluation of various cognitive domains [10–12]. The RBANS takes about 20 min to complete and encompasses 12 subtests organized into five factors: “immediate memory” (F1), “visuospatial/constructional” (F2), “language” (F3), “attention” (F4), and “delayed memory” (F5). Beyond raw scores, it also provides standardized index scores for each factor ($M = 100$, $SD = 15$), derived from the combination of the corresponding subtest scores adjusted for different age groups. The brevity and ease of administration make the RBANS an attractive tool for clinicians and researchers. Additionally, it is composed of alternate forms (also known as parallel forms), which can be extremely useful in a variety of contexts, such as screening and early detection of cognitive impairment, where ongoing monitoring is crucial. Employing alternate forms facilitates repeated assessments over time by mitigating the confounding influence of practice effects.

However, to the best of our knowledge, only one study has compared the VCF administration of the RBANS with its traditional, in-person assessment, with promising but limited results [13]. In this study, no significant difference was found between face-to-face (FtF) and VCF administration. Nevertheless, the analysis relied solely on paired-sample t -tests for the overall sample, without accounting for the effects of age or alternate form used (A vs. B).

Therefore, to face the lack of VCF-based validated tools with alternate forms for cognitive assessment [3, 4, 14, 15], the purpose of the current study was to compare the traditional pen-and-paper version of the RBANS with its VCF-

based counterpart, aiming to ascertain whether there are differences based on the modality employed while addressing the key limitations encountered in the study conducted by Galusha-Glasscock and colleagues [13].

2. Material and Methods

2.1. Participants. A total of 185 participants aged between 20 and 79 years ($M = 46.24$, $SD = 19.63$) took part in the study, of which 58% were women ($N = 108$) and 42% were men ($N = 77$). According to the age ranges proposed by the RBANS scoring system, participants could be categorized in five age groups: 20–39 ($N = 75$; 40 women and 35 men), 40–49 ($N = 25$; 15 women and 10 men), 50–59 ($N = 27$; 16 women and 11 men), 60–69 ($N = 33$; 21 women and 12 men), and 70–79 years ($N = 25$; 16 women and 9 men). Additionally, 172 participants (95.1%) did not have any hearing impairment or had it corrected (e.g., through hearing aids), and 178 participants (96.2%) did not have any visual deficits or had them corrected (e.g., through eyeglasses or contact lenses). Only a few participants reported a history of addiction (1.1%), traumatic brain injury with fainting (3.2%), psychiatric disorders (3.2%), or neurological issues (2.7%). Additionally, 25.9% were taking antidepressants or antipsychotics. No participants were excluded based on these characteristics. Sociodemographic details are provided in Table 1. Sample size was computed a priori using G*Power [16]. Based on a study that employed similar methodologies [17] and the previous study on VCF administration of the RBANS [13], it was estimated that a sample size of at least 24 participants for each of the five age groups was needed for 95% power to detect an F -test effect size of 0.25 using a two-sided alpha level of 0.05.

2.2. Materials. A sociodemographic questionnaire was administered to investigate participants’ background variables (i.e., gender identity, age, level of education, and information on auditory and/or visual deficits).

Forms A and B of the RBANS ([9]; Italian adaptation by [18]) were administered to each participant. The complete structure of the RBANS is reported in Figure 1. Notably, previous validation studies [9, 18] found a significant difference between the two forms in the factor “language” which was due to differences in the subtest “semantic fluency.” In this subtest, participants scored an average of 4 points higher in Form A compared to Form B. Therefore, both Randolph and Ponteri et al. suggested correcting scores in Form B by adding 4 points to make the two forms equivalent.

The Montreal Cognitive Assessment (MoCA) [19] was used as a baseline measure of global cognitive functioning in older adults (aged 65 or above). Permission was granted to use the MoCA in this study.

The Beck Depression Inventory (BDI-II) ([20]; Italian adaptation by [21]) was administered, which consists of 21 multiple-choice items that investigate symptoms of depression experienced in the last 2 weeks (e.g., “sadness”; Cronbach’s $\alpha = 0.90$, current sample).

The Geriatric Depression Scale short version (GDS-15) [22] was used to assess depressive symptoms in older adults.

TABLE 1: Descriptive characteristics of the overall sample (N = 185).

		N	%	Cumulative (%)	M	SD
<i>Categorical variables</i>						
Gender	Women	108	58.4	58.4		
	Men	77	41.6	100		
Level of education	< 8	14	7.6	7.6		
	8–13	77	41.6	41.6		
	> 13	94	50.8	100		
Hearing impairments	No	172	93	93		
	Yes, corrected	4	2.1	95.1		
	Yes, not corrected	9	4.9	100		
Visual impairments	No	55	29.7	29.7		
	Yes, corrected	123	66.5	96.2		
	Yes, not corrected	7	3.8	100		
Addiction (alcohol or drugs)	No	183	98.9	98.9		
	Yes	2	1.1	100		
TBI with loss of consciousness	No	179	96.8	96.8		
	Yes	6	3.2	100		
Diagnosed psychiatric disorders	No	179	96.8	96.8		
	Yes	6	3.2	100		
Strokes/epilepsy/CNS inflammations or infections	No	180	97.3	97.3		
	Yes	5	2.7	100		
Antidepressants or antipsychotic drugs	No	137	74.1	74.1		
	Yes	48	25.9	100		
<i>Continuous variables</i>						
BDI-II		139			2.06	4.88
GDS-15		46			3.50	3.28
STAI-S		185			37.15	9.65
STAI-T		185			41.86	8.99
MoCA		46			24.80	2.93

Abbreviations: Addiction, current or past history of alcohol or drug addiction; BDI-II, Beck Depression Inventory-II (administered when age < 65); CNS, central nervous system; GDS-15, Geriatric Depression Scale short version (administered when age > 64); Level of education, level of education expressed in positively completed years; MoCA, Montreal Cognitive Assessment (administered when age > 64); STAI-S, State-Trait Anxiety Inventory state subscale; STAI-T, State-Trait Anxiety Inventory trait subscale; TBI, traumatic brain injury.

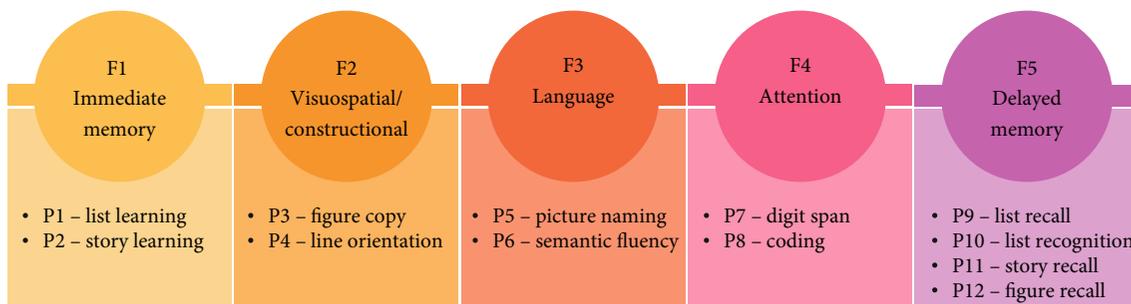


FIGURE 1: Structure of the RBANS.

It consists of 15 yes/no questions (Cronbach’s $\alpha = 0.83$, current sample) that focuses on the cognitive and affective symptoms of depression rather than somatic symptoms, which may be more related to physical health issues common in the elderly.

The State-Trait Anxiety Inventory Form Y (STAI-Y) ([23]; Italian adaptation by [24]) was employed. It is composed of 40 items, of which 20 measure state anxiety (how the participant feels at the time of administration; e.g., “I feel calm”; Cronbach’s $\alpha = 0.90$, current sample) and 20 measure

trait anxiety (how the participant usually feels; e.g., “I feel like a failure”); Cronbach’s $\alpha = 0.83$, current sample).

The short version of the User Experience Questionnaire (UEQ-S) [25] was used to evaluate the usability of the online system. It includes eight pairs of opposite adjectives rated on a 7-point Likert scale, divided into pragmatic (e.g., “complicated/easy”; four items; Cronbach’s $\alpha = 0.82$, current sample) and hedonic (e.g., “boring/exciting”; four items; Cronbach’s $\alpha = 0.86$, current sample) dimensions. Scores below 0 indicate negative evaluations, while scores above 0 indicate positive ones. Results were categorized using benchmarks from Meiners et al. [26] as *excellent*, *good*, *above average*, *below average*, or *bad*, based on comparisons with a diverse sample from various studies.

2.3. Procedure. Participants were recruited on a voluntary basis from the general population to ensure a sample as representative as possible of the general, nonclinical population and were asked to fill out a written informed consent to take part in the research. Data was collected confidentially using an alphanumeric code. The study was conducted in accordance with the ethical principles stated in the Declaration of Helsinki (2013) and approved by the Institutional Review Board of Psychology (IRBP) of the Department of Psychological Science, Health, and Territory at University G. d’Annunzio of Chieti, Pescara (Protocol Number: 20026). Additionally, it adhered to the STROBE guidelines for cross-sectional reporting [27]. The completed checklist is available as supporting information. The study was conducted in two 30-min sessions that were performed at least 1 day and a maximum of 1 week apart. The order of the forms of the RBANS (A and B) and the modality of administration (FtF and VCF) were randomized so that the sample could be equally divided into four conditions: (a) first day: FtF A, second day: VCF B; (b) first day: FtF B, second day: VCF A; (c) first day: VCF A, second day: FtF B; and (d) first day: VCF B, second day: FtF A. Randomization was stratified to balance the four conditions by participants’ gender identity and age using the same age groups employed for the Italian validation of the RBANS [18].

The MoCA and the GDS were administered only to participants aged 65 or above, while the BDI-II was used for participants between 20 and 64 years old. Since it was beyond the scope of the present study, no cut-off was employed for the MoCA scores or other questionnaires, as they served exclusively as control variables to characterize the sample and ensure a thorough understanding of the participants’ baseline conditions. Cronbach’s α was calculated for each self-report questionnaire to assess internal consistency within the current sample.

2.4. Protocol and Adaptation of the RBANS for the Remote Administration. The FtF condition followed the RBANS protocol for in-person assessments. For the VCF administration, the examiner used a 15.6" HP laptop, and participants connected via platforms such as Google Meet or Microsoft Teams. Participants could conduct the remote session from their home environment, having received prior instructions about suitable conditions. Before starting, the administrator

conducted a qualitative audio–video and connection quality check by direct observation and by explicitly asking participants for feedback, categorizing the connection as “poor,” “acceptable,” or “optimal.” Before and during the session, participants were encouraged to report any issues and were invited to pause or stop the session at any moment if connection or hearing/visual problems arose. Any issues occurring during the session were noted by the administrator. If major problems emerged that could drastically impact performance, participants were allowed to finish the tasks, but their data were excluded from the analysis. Notably, they were allowed to choose the device they felt most comfortable using (computer, tablet, or smartphone) for the remote administration in order to better reflect real-world practices, where individuals typically select the device they prefer, as well as to avoid adding potential confounding variables that might arise from forcing a specific instrument. Participants were instructed to prepare two blank sheets of paper and a pencil for the “figure copy” and “figure recall” subtests. In the “coding” subtest, participants read aloud the associated numbers instead of writing them down. After the “figure copy” subtest, participants folded the sheet twice and placed it at a location out of their sight to conceal the drawing. This process was performed in front of the camera, allowing the administrator to verify that participants did not view or access the drawing during the recall phase. Additionally, throughout the session, the administrator monitored to ensure participants did not attempt to retrieve the original figure. At the end of the session, they sent pictures of the sheets via chat, email, or delivered them in person to the examiner.

2.5. Data Analyses. Firstly, a *t*-test was performed to compare Forms A and B on the “semantic fluency” subtest (P6) within the context of FtF administration. If necessary, adjustments to the scores of Form B were made by adding the mean difference found in the FtF administration to the raw scores of P6 in both FtF and VCF administrations.

In the current study, we focused on raw scores to avoid potential confounding issues associated with the RBANS index score system, which are particularly present in the Italian version (for a deeper discussion, see, e.g., [28, 29]). Therefore, factors were calculated by averaging the *Z*-scores of the corresponding subtests. Since these averaged scores were similar but not identical to *Z*-scores, they were then standardized again by converting the factor scores into *Z*-scores to ensure that all factor scores were on a comparable standardized scale. Since subsequent analyses were conducted both on the overall sample and separately for each age group defined by the scoring system, the standardization process was performed twice: first using the mean and standard deviation of the overall sample and then separately for each age group.

Then, linear mixed models (LMMs) in the overall sample were conducted to assess the fixed effect of the modality of administration, form, time (first or second session), and age (included as a continuous variable), as well as the interaction among variables on each factor and subtest of the RBANS. In addition to the evaluation of the overall sample,

TABLE 2: Descriptive statistics and independent sample *t*-tests between Forms A and B (FtF condition only) for each age group on F6 subtest (semantic fluency).

	Form A			Form B			<i>t</i>	df	<i>p</i>	Mean difference	Std. error difference	95% CI	
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>						Lower	Upper
20–39	37	19.54	3.17	38	14.26	4.7	5.69	73	< 0.001	5.28	0.93	3.43	7.13
40–49	14	19.79	3.56	11	13.73	3.93	4.04	23	0.001	6.06	1.50	2.96	9.16
50–59	12	16.00	3.41	15	12.20	4.94	2.26	25	0.033	3.80	1.68	0.34	7.26
60–69	17	18.76	6.20	16	11.44	3.08	4.26	31	< 0.001	7.33	1.72	3.82	10.84
70–79	14	17.00	3.19	11	9.55	2.16	6.64	23	< 0.001	7.45	1.12	5.13	9.78

the analyses to test the fixed effect of modality of administration, form, and time were conducted separately for each age group defined in the FtF norms in order to identify any age-specific differences in performance between administration modalities that might not be apparent in the full sample. Due to the complexity of the LMMs and the relatively low number of participants in each age group, interactions were evaluated in this phase only if found statistically significant in the overall sample. All main effects were retained in the models regardless of statistical significance, with the exception of age, as analyses were already conducted separately by age group.

Moreover, Bayes factors (BFs) were calculated to evaluate the likelihood ratio between the null hypothesis (H0) and the alternative hypothesis (H1, difference between FtF and VCF, with no a priori direction), with H0 in the numerator (BF(0,1)). A BF above 3 indicated moderate evidence for H0, BFs greater than 10 signified strong evidence, and BFs between 0.33 and 3 suggested only anecdotal evidence [30].

Furthermore, intraclass correlation coefficients (ICCs) were computed for each factor using a two-way random effects model, employing a single-rater approach with absolute agreement and with a 95% confidence interval to assess the agreement between the two different modalities of administration (FtF and VCF). Reliability was classified as *poor* below 0.50, *moderate* between 0.50 and 0.75, *good* between 0.75 and 0.90, and *excellent* above 0.90 [31].

Finally, scatterplots were created for each RBANS factor, plotting VCF scores against FtF scores with a best fit line to show the relationship between them. The scores were converted into *T*-scores (positive values only) to make the graphs easier to interpret.

BFs were computed using the “BayesFactor” Package in R. Data from the UEQ-S was analyzed through the freely accessible UEQ Data Analysis Tool (available at <https://www.ueq-online.org/>). All other statistical analyses were conducted using SPSS 25.

3. Results

Data were close to the normal distribution for all the subtests except for the “line orientation” (P4) and the “picture naming” (P5) subtests, which showed a left-skewed distribution. Notably, in P5, in which the maximum score was 10, most of the sample reached the maximum score with very small variability ($M = 9.88$, $SD = 0.32$). On the other side, scores of P4, in which the maximum score was 20, were more heterogeneous ($M = 16.19$, $SD = 3.30$).

In the overall sample, a significant difference was observed between FtF Forms A and B of the RBANS in P6 (“semantic fluency”). Participants scored approximately 6 points lower on Form B compared to Form A ($t(183) = 9.29$; $p < 0.001$). As shown in Table 2, the mean differences varied by age group. To address the heterogeneity of mean differences across age groups, P6 scores were adjusted accordingly. Specifically, in analyses involving the overall sample, the mean difference observed across the entire sample was applied, adding 6 points to all individual scores on P6 Form B. For subgroup analyses, scores were adjusted separately for each group by adding the corresponding mean difference to Form B.

In the overall sample, tests of fixed effects using LMMs revealed that age was significant for each factor and each subtest of the RBANS (p values were ≤ 0.003), in which higher age was associated with lower scores. Furthermore, the interaction between time and form was significant in the “digit span” subtest (P7) ($F(1,177) = 4.94$; $p = 0.027$). However, the interaction was no longer significant when conducting pairwise comparisons. Additionally, in “delayed memory” factor (F5) there was a significant effect of time ($F(1,177) = 8.320$; $p = 0.004$) with a mean difference (ΔM) of 0.18, in which scores were significantly higher in the second session (Time 2) compared to the first one (Time 1). Accordingly, time was significant in two out of four F5’s subtests, “list recall” (P9) ($F(1,177) = 17.71$; $p < 0.001$; $\Delta M = 0.13$) and “list recognition” (P10) ($F(1,177) = 8.96$; $p = 0.003$; $\Delta M = 0.44$), with participants obtaining higher scores at Time 2. The interaction between time and age was significant for in the “list recall” subtest (P9) ($F(1,177) = 4.34$; $p = 0.037$), in which scored were higher at Time 2, especially for younger adults where the discrepancy was more pronounced compared to older participants. Moreover, a significant interaction between form and age emerged in “story recall” (P11) ($F(1,177) = 3.99$; $p = 0.047$). Finally, form was significant in P10 ($F(1,177) = 4.59$; $p = 0.033$; $\Delta M = 0.29$), in which Form B showed significantly higher scores compared to Form A. The complete results of LMMs in the overall sample for each factor and for each subtest can be found in Tables 3 and 4, respectively. No additional effects were statistically significant.

In summary, while age and time showed significant effects across various factors and subtests of the RBANS in the overall sample ($p < 0.05$), there was consistently no observed effect of modality of administration ($p > 0.05$).

TABLE 3: Tests of fixed effects in the overall sample ($N = 185$) for each factor of the RBANS.

	F1	F2	F3	F4	F5
Form	2.122 (0.147)	0.230 (0.632)	3.014 (0.084)	0.474 (0.492)	1.312 (0.253)
Modality	0.829 (0.364)	0.002 (0.969)	0.280 (0.598)	0.045 (0.832)	0.035 (0.853)
Time	0.545 (0.461)	0.389 (0.533)	2.276 (0.133)	0.698 (0.405)	8.320 (0.004)
Age	65.124 (< 0.001)	115.126 (< 0.001)	9.029 (0.003)	100.753 (< 0.001)	63.300 (< 0.001)
Form * modality	0.188 (0.665)	0.070 (0.791)	1.332 (0.250)	0.463 (0.497)	0.171 (0.680)
Form * time	0.124 (0.725)	0.300 (0.585)	0.030 (0.862)	0.663 (0.416)	0.001 (0.980)
Form * age	0.554 (0.458)	0.387 (0.534)	0.738 (0.392)	0.729 (0.394)	0.012 (0.915)
Modality * time	2.555 (0.135)	0.570 (0.451)	0.568 (0.452)	0.959 (0.329)	1.582 (0.210)
Modality * age	1.016 (0.315)	0.461 (0.498)	0.043 (0.835)	0.074 (0.786)	0.002 (0.966)
Time * age	0.448 (0.504)	0.390 (0.533)	3.086 (0.081)	0.000 (0.990)	3.291 (0.071)
Form * modality * time	0.210 (0.647)	0.205 (0.651)	1.604 (0.207)	0.616 (0.434)	0.001 (0.980)
Form * modality * age	0.658 (0.418)	0.087 (0.769)	0.347 (0.556)	0.964 (0.328)	0.003 (0.959)
Form * time * age	0.297 (0.587)	0.044 (0.833)	0.113 (0.737)	3.109 (0.080)	0.007 (0.934)
Modality * time * age	1.342 (0.248)	0.524 (0.470)	0.399 (0.528)	0.177 (0.675)	1.991 (0.160)
Form * modality * time * age	0.238 (0.626)	0.361 (0.549)	0.663 (0.416)	0.247 (0.620)	0.462 (0.498)

Note: Values indicate F -statistic value with $F(1,177)$, followed by its associated p value in parenthesis. Significant effects ($p < 0.05$) are reported in bold. Abbreviations: F1, immediate memory; F2, visuospatial/construction; F3, language; F4, attention; F5, delayed memory.

When considering age groups separately, the only significant effect on the RBANS factors was time, which was significant for the 20–39 group in F5 (“delayed memory”) ($F(1, 72) = 11.82$; $p = 0.001$), favoring Time 2 with a mean difference of 0.337, and for the 60–69 group in F4 (“attention”) ($F(1, 30) = 5.12$; $p = 0.031$), favoring Time 2 with a 0.29 mean difference.

Taking into account subtests, the effect of time was significant for two F5 subtests only in age groups 20–39 (P9: $F(1, 72) = 38.42$; $p < 0.001$; $\Delta M = 0.54$; P10: $F(1, 72) = 18.70$; $p < 0.001$; $\Delta M = 0.61$) and 60–69 (P9: $F(1, 30) = 7.31$; $p = 0.011$; $\Delta M = 0.50$; P10: $F(1, 30) = 16.72$; $p < 0.001$; $\Delta M = 0.58$), in which scores were higher at Time 2 compared to Time 1. Additionally, it was significant for the 20–39 group in P1 ($F(1, 72) = 4.29$; $p = 0.042$; $\Delta M = 0.07$), for the 40–49 group in P4 ($F(1, 22) = 4.80$; $p = 0.039$; $\Delta M = 0.45$), and in P8 for the 20–39 age group ($F(1, 72) = 8.25$; $p = 0.005$; $\Delta M = 0.23$) and the 60–69 age group ($F(1, 30) = 9.41$; $p = 0.005$; $\Delta M = 0.29$). Finally, a significant effect of form was found in the “story learning” subtest for the 50–59 age group (P2) ($F(1, 25) = 4.39$; $p = 0.047$), in which scores were significantly higher in Form B compared to Form A with a mean difference (ΔM) of 0.46 standard deviations; in the “line orientation” subtest for the age groups 20–39 (P4) ($F(1, 72) = 7.40$; $p = 0.008$; $\Delta M = 0.31$) and 60–69 ($F(1, 30) = 5.07$; $p = 0.032$; $\Delta M = 0.32$), in which participants in both age groups obtained significantly higher scores in Form A; in the “picture naming” subtest for the 50–59 age group (P5) ($F(1, 24) = 5.38$; $p = 0.029$; $\Delta M = 0.61$), in which scores were significantly higher in Form B; and in the “list recognition” subtest for the 20–39 age group (P10) ($F(1, 72) = 8.81$; $p = 0.004$; $\Delta M = 0.41$), in which Form B showed significantly higher scores compared to Form A. No other significant effects were observed. With respect to the

scores of the “picture naming” subtest (P5) obtained by the 50–59 age group, in which a significant effect of form was found, it emerged that the maximum score was reached 81% of times for Form A and 100% of times for Form B. More details about the results of LMMs conducted separately for each age group are reported as supporting information (please see Tables S1 and S2).

Overall, analyses on age groups showed that the effect of the modality of administration was never significant ($p > 0.05$). The effect of time and form emerged only in some subtests and exclusively in specific age groups.

Considering the modality of administration, results from the Bayesian analysis supported the null model for each factor and each subtest: BFs were between 5.60 and 8.70 for factors and 3.99 and 8.61 for subtests. Similarly, BFs computed for the interaction between modality and form supported the null model. Apart from “language” (F3) and “picture naming” (P5), which were 1.78 and 2.85, respectively, all the other BFs were greater than 3 (between 11.19 and 96.43 for factors and between 7.03 and 169.49 for subtests). Finally, BFs for form were in favor of the alternative model (H1) for F3 (BF = 0.39), P5 (BF = 0.10), and “list recognition” (P10) (BF = 0.75); all the other BFs computed for form supported the null model: apart from the remaining factors and for the “story learning” (P2) and “line orientation” (P4) subtests, in which BFs were less than 3 (but greater than 1), all the other BFs ranged between 5.66 and 8.70. Thus, BFs showed moderate evidence for the null model when considering the modality of administration, anecdotal through very strong evidence based on the interaction between the factors, while anecdotal through moderate evidence was found regarding form, apart from F3, P5, and P10, in which anecdotal through moderate evidence in favor of the alternative model was found.

TABLE 4: Tests of fixed effect for each subtest of the RBANS in the overall sample (N = 185).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
Form	0.834 (0.362)	1.462 (0.228)	0.467 (0.495)	1.46 (0.229)	3.119 (0.079)	0.992 (0.321)	0.067 (0.797)	0.828 (0.364)	1.618 (0.205)	4.595 (0.033)	1.804 (0.181)	0.258 (0.612)
Modality	0.325 (0.569)	3.816 (0.052)	0.12 (0.729)	0.115 (0.735)	2.538 (0.112)	0.75 (0.388)	0.282 (0.596)	0.254 (0.615)	1.903 (0.170)	0.056 (0.814)	0.506 (0.478)	0.35 (0.355)
Time	2.299 (0.131)	0.243 (0.623)	0.088 (0.767)	1.169 (0.281)	0.235 (0.628)	2.469 (0.118)	0.004 (0.948)	2.356 (0.127)	17.714 (< 0.001)	8.956 (0.003)	0.222 (0.638)	0.094 (0.760)
Age	79.523 (> 0.001)	24.752 (> 0.001)	59.551 (< 0.001)	52.028 (< 0.001)	9.371 (0.003)	18.727 (> 0.001)	14.245 (> 0.001)	158.642 (> 0.001)	21.927 (> 0.001)	13.452 (< 0.001)	26.869 (> 0.001)	69.301 (> 0.001)
Form * modality	0.135 (0.714)	1.041 (0.309)	0.29 (0.591)	0.029 (0.865)	2.289 (0.132)	0.205 (0.652)	1.017 (0.314)	0.002 (0.963)	0.005 (0.945)	0.161 (0.689)	0.093 (0.761)	1.285 (0.259)
Form * time	0.013 (0.910)	0.214 (0.644)	0.136 (0.713)	0.155 (0.694)	0.684 (0.409)	1.166 (0.282)	4.955 (0.027)	1.366 (0.244)	0.023 (0.880)	0.975 (0.325)	0.126 (0.724)	2.053 (0.154)
Form * age	0.814 (0.368)	0.026 (0.872)	0.867 (0.353)	0.004 (0.951)	0.415 (0.521)	0.277 (0.600)	0.164 (0.686)	0.962 (0.328)	1.289 (0.258)	0.806 (0.371)	3.989 (0.047)	0.047 (0.829)
Modality * time	0.502 (0.480)	3.139 (0.078)	2.557 (0.112)	0.306 (0.581)	0.846 (0.359)	3.032 (0.083)	1.936 (0.166)	0.001 (0.996)	0.534 (0.466)	0.157 (0.692)	2.431 (0.121)	1.936 (0.166)
Modality * age	0.399 (0.528)	0.378 (0.539)	0.485 (0.487)	0.11 (0.740)	0.808 (0.370)	0.339 (0.561)	1.057 (0.305)	1.563 (0.213)	1.55 (0.215)	0.052 (0.820)	0.105 (0.746)	0.913 (0.341)
Time * age	0.553 (0.458)	0.047 (0.829)	0.331 (0.566)	0.129 (0.720)	1.184 (0.278)	2.528 (0.114)	0.001 (0.995)	0.001 (0.990)	4.434 (0.037)	0.983 (0.323)	0.514 (0.474)	0.242 (0.623)
Form * modality * time	0.74 (0.391)	2.514 (0.115)	0.578 (0.448)	0.001 (0.981)	2.836 (0.094)	0.056 (0.813)	1.412 (0.236)	0.266 (0.607)	0.051 (0.821)	0.002 (0.962)	0.001 (0.975)	0.033 (0.856)
Form * modality * age	0.008 (0.93)	1.896 (0.170)	0.015 (0.902)	0.082 (0.775)	1.491 (0.224)	0.027 (0.869)	1.226 (0.270)	0.122 (0.727)	0.08 (0.777)	0.031 (0.862)	0.392 (0.532)	0.813 (0.369)
Form * time * age	0.069 (0.794)	0.408 (0.524)	0.002 (0.963)	0.116 (0.734)	1.867 (0.173)	0.909 (0.342)	2.825 (0.095)	0.992 (0.321)	0.029 (0.864)	0.678 (0.412)	0.330 (0.567)	1.829 (0.178)
Modality * time * age	0.006 (0.941)	3.193 (0.076)	3.752 (0.054)	0.877 (0.350)	1.447 (0.231)	3.402 (0.067)	1.197 (0.275)	0.349 (0.555)	1.015 (0.315)	0.127 (0.722)	1.447 (0.231)	1.015 (0.315)
Form * modality * time * age	0.242 (0.623)	1.543 (0.216)	0.095 (0.759)	0.302 (0.583)	2.067 (0.152)	0.053 (0.818)	1.475 (0.226)	1.276 (0.260)	2.114 (0.148)	0.139 (0.709)	5.162 (0.808)	0.416 (0.520)

Note: Values indicate *F*-statistic value with *F*(1,177), followed by its associated *p* value in parenthesis. Significant effects (*p* < 0.05) are reported in bold.

Abbreviations: P1, list memory; P2, story memory; P3, figure copy; P4, line orientation; P5, picture naming; P6, semantic fluency; P7, digit span; P8, coding; P9, list recall; P10, list recognition; P11, story recall; P12, figure recall.

ICCs were 0.65 for F1, 0.73 for F2, 0.17 for F3, 0.77 for F4, and 0.66 for F5. Overall, factors indicated moderate agreement between the two modalities of administration, except for F3, in which it was poor. Figure 2 presents scatter plots with regression lines illustrating the relationship between FtF and VCF modalities of administration. Points above the best fit line indicate that the observed VCF scores were higher than what the best fit line predicted for the corresponding FtF scores, while points below the best fit line indicate that the observed VCF scores were lower than what the best fit line predicted for the corresponding FtF scores. The points in the scatter plots were fairly symmetrical, indicating that the number of participants who performed better in the VCF modality compared to the FtF modality was approximately equal to the number who performed better in the FtF modality compared to the VCF modality. The slope of the best fit line was acceptable (close to 1), as well as the effect size ($R^2 \geq 0.25$), across most factors, except for F3.

Finally, results from the UEQ-S showed that none of the participants rated the VCF-based administration of the RBANS below 0. Comparing current ratings with the available benchmarks, the VCF administration was rated as *excellent* ($M = 1.67$). Particularly, the pragmatic dimension was rated as *good* ($M = 1.51$), while the hedonic dimension was rated as *excellent* ($M = 1.82$).

4. Discussion

The present study was aimed at evaluating the consistency of the RBANS [9, 18] scores and their interpretation across different administration modalities, namely, FtF and VCF. Specifically, participants underwent RBANS testing twice: once in person, utilizing standard pen-and-paper administration, and once remotely via VCF, using alternate forms to mitigate practice effects. In line with previous research [13], no significant differences were observed based on the modality of administration, a result that was reinforced by the Bayesian analysis. Importantly, while a previous study on the remote administration of the RBANS [13] found no differences between FtF- and VCF-based administrations, it did not analyze the effect of age, nor consider age subgroups separately. Our study filled this gap, revealing no significant differences in administration mode, even when controlling for age and evaluating age groups separately. Overall, findings indicated that older adults can be reliably assessed remotely using the RBANS VCF adaptation, which not only enhances the inclusivity of the cognitive assessment but also demonstrates the adaptability and robustness of remote assessment protocols in accommodating diverse populations.

However, discrepancies emerged in certain subtests when comparing the alternate forms of the RBANS. Notably, although a significant difference was expected between Forms A and B in the subtest “semantic fluency” [18], in the current study, it exceeded prior expectations and was particularly evident in the 60–69 and 70–79 age groups, where it reached approximately 7 points. Notably, while Form A refers to a broad category (“fruits and vegetables”), Form B pertains to a semantic subcategory (“animals found

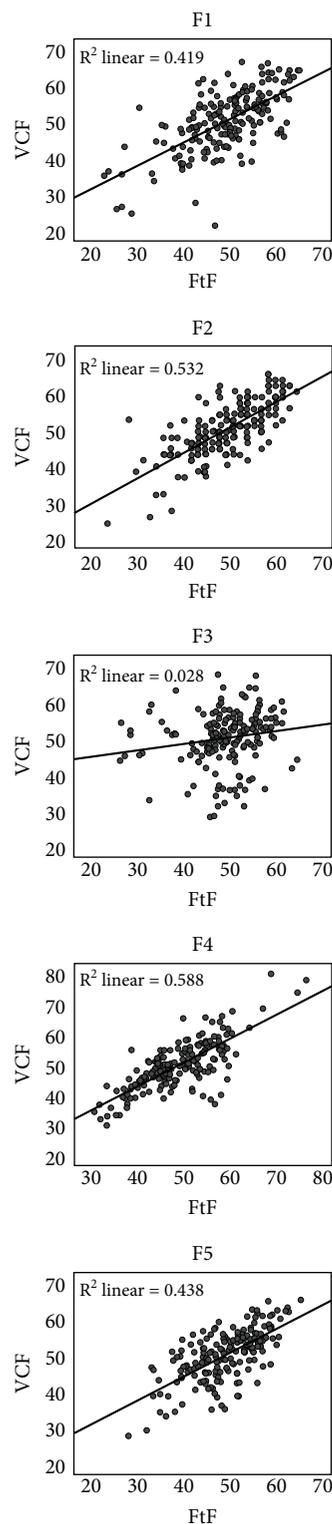


FIGURE 2: Scatter plots of each factor of the RBANS, with each point representing an observation and the regression line indicating the best fit line for predicting VCF from FtF scores.

in a zoo”), which might demand more cognitive effort from the participants. Furthermore, some discrepancies were observed in other RBANS subtests within age groups, showing heterogeneous patterns, predominantly in the 20–39,

50–59, and 60–69 age groups. Moreover, a significant effect of time emerged in the “delayed memory” factor (F5), and more specifically in two of its subtests (“list recall” and “list recognition”), showing that scores were higher the second time these subtests were completed. This result is not surprising, as the standardized instructions for administering the RBANS require that participants are not informed during the “immediate memory” subtests that they will later be asked to recall the items, leading to more difficulty. By the second session, participants were aware of the recall task and likely exerted more effort to remember the items. Additionally, it seems unlikely that significant difference depended on the selected time range, as alternate forms of the RBANS may technically be administered even on the same day. We chose the current interval to align with previous studies, including the original and Italian validations [9, 13, 18], but learning effect are not expected to influence performance based on the number of days between sessions even within this short timeframe.

Overall, these findings point to some limitations associated with the repeated administration of the RBANS, even when alternate forms are used. On one side, variations between the alternate forms have been observed in certain subtests, which may raise concerns about test equivalence and the reliability of standardized administration. Furthermore, the presence of a ceiling effect in one of the “language” subtests (“picture naming”) may have further confounded result interpretation [32]. Finally, repeated testing may lead to practice effects, which can make direct comparisons of some scores—such as delayed memory—unreliable [33].

To mitigate these issues, it is advisable to inform participants about the upcoming memory recall task during the tasks related to immediate memory and to ensure thorough validation of alternate forms. Consistently, updated standardized index scores may be necessary for each alternate form of the RBANS using a broader normative sample [28].

To summarize, the findings of this study support the use of remote RBANS administration, indicating that test scores are consistent and comparable across modalities. However, they also highlight significant issues that underscore the importance of ongoing validation and refinement of remote neuropsychological assessment protocols. Indeed, as no significant effects emerged for the modality of administration, effect sizes were not further examined for variables such as test form, which were beyond the main focus of the study. However, it is worth noting that in the context of a normative scoring system, even small differences may have important implications for clinical interpretation and should be further explored in future research, specifically to determine whether normative adjustments may be necessary. While these challenges are particularly evident in the RBANS, the insights and recommendations drawn from this research are broadly applicable to the development and refinement of all cognitive assessment tools. Clinicians and researchers must be vigilant regarding the potential limitations in reliability and validity of the assessment tools employed. Addressing these issues is crucial to ensure the overall reliability, validity, and utility of remote assessment tools in both clinical practice and research settings. Indeed, by addressing these challenges, the precision and effectiveness

of cognitive assessment can be enhanced, ultimately improving diagnostic accuracy and treatment outcomes for individuals with cognitive impairment.

Despite the inherent limitations of the RBANS, it is remarkable that its adaptation for remote assessment can be utilized effectively, which may imply a significant advancement in the accessibility and reliability of neuropsychological cognitive assessments [2]. Notably, the results suggested that participants rated the VCF administration of the RBANS positively across all dimensions, indicating that the VCF format was well received, providing a satisfactory user experience in terms of both functionality and enjoyability. Moreover, it is interesting to note its applicability among older adults aged 65 and over, a demographic group that may be less experienced with technology and potentially more susceptible to technological influences on test results.

4.1. Limitations and Future Directions. Despite the valuable insights gained from the present study, some limitations warrant consideration. Firstly, the sample size, while sufficient for conducting statistical analyses, was relatively small within each age group. This limitation might indicate issues in detecting statistical differences for small effect sizes; therefore, caution is needed when interpreting results disaggregated by age groups. Moreover, the generalizability of findings may be compromised by the characteristics of the sample. Participants who volunteered for the study may be more inclined to use technology, potentially biasing the sample towards individuals comfortable with remote cognitive assessment. For example, older adults with limited technological literacy or an aversion to technology may have self-selected out of participation, limiting the representation of this demographic in the study. Additionally, it was not possible to compare participants’ performance based on the type of device used during the VCF administration. Future studies should investigate whether specific device characteristics, such as screen size or input modality, may influence cognitive performance. Similarly, it would be important to further explore the role of individual differences in shaping performance across administration modalities, including variables such as educational level, digital literacy, autonomy in the use of technological devices, mood, and other psychosocial factors.

To minimize variability in test performance, future research should focus on being specifically designed to assess differences across test forms and determine whether effects exist that may justify updating and developing normative scores for alternate forms of the RBANS so that they can be accurately compared, taking effect sizes and their potential clinical implications into account. This is crucial also for evaluating the remote administration of the RBANS, as combining different forms and administration modalities can obscure results, making it difficult to accurately assess the impact of the modality of administration when alternate forms are used.

5. Conclusion

In conclusion, this study examined potential differences between the remote (through VCF) and the traditional FtF administration of the RBANS. Findings suggested that the same norms that are used for the FtF administration of the

RBANS may also be applied when the RBANS is administered remotely through VCF. The possibility of remotely administering standardized tests with alternate forms, such as the RBANS, is crucial in clinical practice as it enhances access to neuropsychological assessment services, improves monitoring of cognitive function over time, facilitates early detection of cognitive changes, and enhances patient care.

Nevertheless, continued validation and refinement of remote assessment protocols, including the RBANS, are necessary to ensure their reliability, validity, and utility across diverse populations and clinical contexts. While challenges remain, the promising results of this study underscore the potential of remote neuropsychological assessment as a valuable tool in the assessment and monitoring of cognitive function, particularly considering the evolving healthcare needs and technological advancements.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Ethics Statement

The study was conducted in accordance with the ethical principles stated in the Declaration of Helsinki (2013) and approved by the Institutional Review Board of Psychology (IRBP) of the Department of Psychological Science, Health, and Territory at University G. d'Annunzio of Chieti, Pescara.

Consent

All participants provided written informed consent prior to enrollment in the study.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

No funding was received for this manuscript.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. (*Supporting Information*) The following supporting information is provided in the "Supporting Information" file: Table S1: Tests of fixed effects for each RBANS factor, separated by age group. Table S2: Tests of fixed effects for each RBANS subtest, separated by age group. STROBE Checklist: Completed STROBE checklist for cross-sectional studies.

References

[1] A. A. Tahami Monfared, Y. Stern, S. Doogan, M. Irizarry, and Q. Zhang, "Understanding Barriers Along the Patient Journey in Alzheimer's Disease Using Social Media Data," *Neurology*

and *Therapy* 12, no. 3 (2023): 899–918, <https://doi.org/10.1007/s40120-023-00472-x>.

[2] T. Parsons and T. Duffield, "Paradigm Shift Toward Digital Neuropsychology and High-Dimensional Neuropsychological Assessments: Review," *Journal of Medical Internet Research* 22, no. 12 (2020): e23777, <https://doi.org/10.2196/23777>.

[3] D. Binng, M. Splonskowski, and C. Jacova, "Distance Assessment for Detecting Cognitive Impairment in Older Adults: A Systematic Review of Psychometric Evidence," *Dementia and Geriatric Cognitive Disorders* 49, no. 5 (2021): 456–470, <https://doi.org/10.1159/000511945>.

[4] T. W. Brearly, R. D. Shura, S. L. Martindale, et al., "Neuropsychological Test Administration by Videoconference: A Systematic Review and Meta-Analysis," *Neuropsychology Review* 27, no. 2 (2017): 174–186, <https://doi.org/10.1007/s11065-017-9349-1>.

[5] A. D. Mishkin, J. S. Zabinski, G. Holt, and P. S. Appelbaum, "Ensuring Privacy in Telemedicine: Ethical and Clinical Challenges," *Journal of Telemedicine and Telecare* 29, no. 3 (2023): 217–221, <https://doi.org/10.1177/1357633X221134952>.

[6] E. Şahin, B. G. Yavuz Veizi, and M. I. Naharci, "Telemedicine Interventions for Older Adults: A Systematic Review," *Journal of Telemedicine and Telecare* 30, no. 2 (2024): 305–319, <https://doi.org/10.1177/1357633X211058340>.

[7] T. Schroeder, L. Dodds, A. Georgiou, H. Gewald, and J. Siette, "Older Adults and New Technology: Mapping Review of the Factors Associated With Older Adults' Intention to Adopt Digital Technologies," *JMIR Aging* 6, no. 1 (2023): e44564, <https://doi.org/10.2196/44564>.

[8] B. León-Salas, Y. González-Hernández, D. Infante-Ventura, et al., "Telemedicine for Neurological Diseases: A Systematic Review and Meta-Analysis," *European Journal of Neurology* 30, no. 1 (2023): 241–254, <https://doi.org/10.1111/ene.15599>.

[9] C. Randolph, *Repeatable Battery for the Assessment of Neuropsychological Status (RBANS)* (Psychological Corporation San Antonio, 1998).

[10] K. Duff, H. J. D. Clark, S. E. O'Bryant, J. W. Mold, R. B. Schiffer, and P. B. Sutker, "Utility of the RBANS in Detecting Cognitive Impairment Associated With Alzheimer's Disease: Sensitivity, Specificity, and Positive and Negative Predictive Powers," *Archives of Clinical Neuropsychology* 23, no. 5 (2008): 603–612, <https://doi.org/10.1016/j.acn.2008.06.004>.

[11] K. Duff, V. L. Hobson, L. J. Beglinger, and S. E. O'Bryant, "Diagnostic Accuracy of the RBANS in Mild Cognitive Impairment: Limitations on Assessing Milder Impairments," *Archives of Clinical Neuropsychology* 25, no. 5 (2010): 429–441, <https://doi.org/10.1093/arclin/acq045>.

[12] R. D. Shura, T. W. Brearly, J. A. Rowland, S. L. Martindale, H. M. Miskey, and K. Duff, "RBANS Validity Indices: A Systematic Review and Meta-Analysis," *Neuropsychology Review* 28, no. 3 (2018): 269–284, <https://doi.org/10.1007/s11065-018-9377-5>.

[13] J. M. Galusha-Glasscock, D. K. Horton, M. F. Weiner, and C. M. Cullum, "Video Teleconference Administration of the Repeatable Battery for the Assessment of Neuropsychological Status," *Archives of Clinical Neuropsychology* 31, no. 1 (2016): 8–11, <https://doi.org/10.1093/arclin/acv058>.

[14] E. Salvadori, L. Pantoni, and S. I. di NeuroGeriatrics, "Teleneuropsychology for Vascular Cognitive Impairment: Which Tools Do We Have?," *Cerebral Circulation-Cognition and Behavior* 5 (2023): 100173, <https://doi.org/10.1016/j.cccb.2023.100173>.

- [15] E. Zanin, E. N. Aiello, L. Diana, et al., “Tele-Neuropsychological Assessment Tools in Italy: A Systematic Review on Psychometric Properties and Usability,” *Neurological Sciences* 43, no. 1 (2022): 125–138, <https://doi.org/10.1007/s10072-021-05719-9>.
- [16] F. Faul, E. Erdfelder, A.-G. Lang, and A. Buchner, “G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences,” in *Behavior Research Methods* 39, no. 2 (Psychonomic Society, 2007), 175–191, <https://doi.org/10.3758/BF03193146>.
- [17] A. D. Rebchuk, H. M. Deptuck, Z. R. O’Neill, D. S. Fawcett, N. D. Silverberg, and T. S. Field, “Validation of a Novel Telehealth Administration Protocol for the NIH Toolbox-Cognition Battery,” *Telemedicine and E-Health* 25, no. 3 (2019): 237–242, <https://doi.org/10.1089/tmj.2018.0023>.
- [18] M. Ponteri, R. Pioli, A. Padovani, S. Tunesi, and G. De Girolamo, *RBANS Repeatable Battery for the Assessment of Neuropsychological Status. Edizione Italiana* (Giunti OS, 2007).
- [19] Z. S. Nasreddine, N. A. Phillips, V. Bédirian, et al., “The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool for Mild Cognitive Impairment,” *Journal of the American Geriatrics Society* 53, no. 4 (2005): 695–699, <https://doi.org/10.1111/j.1532-5415.2005.53221.x>.
- [20] A. T. Beck, R. A. Steer, and G. Brown, “Beck Depression Inventory–II,” *Corsini Encyclopedia Psychology* 1, no. 1 (1996): 210.
- [21] M. Ghisi, G. B. Flebus, A. Montano, E. Sanavio, and C. Sica, *Beck Depression Inventory-II. Manuale Italiano* (Giunti Organizzazioni Speciali, 2006).
- [22] J. I. Sheikh and J. A. Yesavage, “Geriatric Depression Scale (GDS): Recent Evidence and Development of a Shorter Version,” in *Clinical Gerontology* (Routledge, 1986), 165–173.
- [23] C. D. Spielberger, R. Goruch, R. Lushene, P. Vagg, and G. Jacobs, *Manual for the State-Trait Inventory STAI (Form Y)* (Mind Garden, 1983).
- [24] L. Pedrabissi and M. Santinello, “Verifica della validità dello STAI forma Y di Spielberger,” in *Giunti Organizzazioni Speciali* (Organizzazioni Speciali Firenze, 1989), 11–14.
- [25] M. Schrepp, A. Hinderks, and J. Thomaschewski, “Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S),” *International Journal of Interactive Multimedia and Artificial Intelligence* 4, no. 6 (2017): 103–108, <https://doi.org/10.9781/ijimai.2017.09.001>.
- [26] A. L. Meiners, M. Schrepp, A. Hinderks, and J. Thomaschewski, “A Benchmark for the UEQ+ Framework: Construction of a Simple Tool to Quickly Interpret UEQ+ KPIs,” 2024): <https://doi.org/10.9781/ijimai.2023.05.003>.
- [27] E. von Elm, D. G. Altman, M. Egger, et al., “The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies,” *PLoS Medicine* 4, no. 10 (2007): e296, <https://doi.org/10.1371/journal.pmed.0040296>.
- [28] E. Chipi, D. Fruttini, N. Salvadori, et al., “Repeatable Battery for the Assessment of Neuropsychological Status: Italian Normative Data for Older Adults,” *Archives of Clinical Neuropsychology* 38, no. 1 (2023): 72–79, <https://doi.org/10.1093/arclin/acac061>.
- [29] T. Galt, D. De Ridder, P. Manning, and N. Cutfield, “RBANS Form Equivalence in Specific English Language Regions,” *New Zealand Journal of Psychology* 50, no. 3 (2021): 55–60.
- [30] C. Keyzers, V. Gazzola, and E. J. Wagenmakers, “Using Bayes Factor Hypothesis Testing in Neuroscience to Establish Evidence of Absence,” *Nature Neuroscience* 23, no. 7 (2020): 788–799, <https://doi.org/10.1038/s41593-020-0660-4>.
- [31] T. K. Koo and M. Y. Li, “A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research,” *Journal of Chiropractic Medicine* 15, no. 2 (2016): 155–163, <https://doi.org/10.1016/j.jcm.2016.02.012>.
- [32] B. Uttl, “Measurement of Individual Differences,” *Psychological Science* 16, no. 6 (2005): 460–467, <https://doi.org/10.1111/j.0956-7976.2005.01557.x>.
- [33] T. E. Goldberg, P. D. Harvey, K. A. Wesnes, P. J. Snyder, and L. S. Schneider, “Practice Effects due to Serial Cognitive Assessment: Implications for Preclinical Alzheimer’s Disease Randomized Controlled Trials,” *Alzheimer’s & Dementia: Diagnosis, Assessment & Disease Monitoring* 1, no. 1 (2015): 103–111, <https://doi.org/10.1016/j.dadm.2014.11.003>.