

Marquette University

e-Publications@Marquette

Psychology Faculty Research and Publications

Psychology, Department of

2017

Invariant Two-Component Structure of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS)

Elisabeth M. Vogt

Marquette University, elisabeth.vogt@marquette.edu

Gregory D. Prichett

Gunderson Health System

James Hoelzle

Marquette University, james.hoelzle@marquette.edu

Follow this and additional works at: https://epublications.marquette.edu/psych_fac



Part of the [Psychology Commons](#)

Recommended Citation

Vogt, Elisabeth M.; Prichett, Gregory D.; and Hoelzle, James, "Invariant Two-Component Structure of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS)" (2017). *Psychology Faculty Research and Publications*. 267.

https://epublications.marquette.edu/psych_fac/267

Marquette University

e-Publications@Marquette

Psychology Faculty Research and Publications/College of Arts and Sciences

This paper is NOT THE PUBLISHED VERSION.

Access the published version at the link in the citation below.

Proceedings of the 23rd International Conference on Software Engineering. ICSE 2001, (May 19, 2001).
[DOI](#). This article is © The Institute of Electrical and Electronics Engineers and permission has been granted for this version to appear in [e-Publications@Marquette](#). The Institute of Electrical and Electronics Engineers does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from The Institute of Electrical and Electronics Engineers.

Invariant Two-component Structure of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS).

Elisabeth M. Vogt

Department of Psychology, Marquette University, Milwaukee, WI

Gregory D. Prichett

Neuropsychology, Gundersen Health System, La Crosse, WI

James B. Hoelzle

Department of Psychology, Marquette University, Milwaukee, WI

Abstract

The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, 1998, 2012) is a brief neurocognitive instrument used to evaluate cognitive functioning in clinical settings. Prior investigations of the factor structure have revealed subtle differences across samples. It was

hypothesized that these differences are primarily the result of methodological decisions made by researchers. The present study utilized empirically supported extraction criteria (parallel analysis; minimum average partial procedure) and uniformly investigated 5 samples. RBANS data from 4 previously published studies (Carlozzi, Horner, Yang, & Tilley, 2008; Duff, Hobson, Beglinger, & O'Bryant, 2010; Duff et al., 2006; Wilde, 2006) were reanalyzed, and a new clinical sample was obtained from the Gundersen Health System Memory Center. The congruence of factor structures was investigated by conducting orthogonal vector matrix comparisons (Barrett, 2005), and a robust 2-factor structure reliably emerged across samples. The invariant RBANS 2-factor structure primarily emphasized memory and visuospatial functioning. This finding offered further support for a 2-factor RBANS structure identified in previous studies and additionally provided empirical documentation of replication across diverse samples. Due to the expansive use of the RBANS, this psychometric knowledge has significant clinical implications. It should facilitate accurate interpretation of test data and allow clinicians to make more informed decisions regarding whether the instrument is appropriate to use in various clinical settings.

Keywords

Assessment; neuropsychology; principal components analysis; psychometrics; Repeatable Battery for the Assessment of Neuropsychological Status (RBANS)

Introduction

The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, [58]; RBANS Update, Randolph, [59]) was developed to address a need for brief assessment measures that are sensitive to impairment in multiple cognitive domains. Consideration of cognitive theory and neuropsychological functioning guided development of the battery (Schoenberg & Scott, [61]). While the RBANS subtests parallel frequently utilized and well-validated neuropsychological measures (Camara, Nathan, & Puente, [9]), they include fewer items and therefore result in quicker administration. The RBANS evaluates a broad range of cognitive abilities and has shown utility in a variety of clinical settings (Aupperle, Beatty, Shelton, & Gontkovsky, [1]; Beatty et al., [6]; Larson, Kirschner, Bode, Heinemann, & Goodman, 2005; McKay, Casey, Wertheimer, & Fichtenberg, [50]; Wilk et al., [5]).

Since publication of the RBANS in 1998, multiple studies have evaluated the reliability, validity, and clinical utility of the measure. Although this measure was originally developed for dementia evaluations, clinicians have utilized the RBANS as a key aspect of assessment across multiple clinical populations such as those presenting with Parkinson disease (Beatty et al., [6]), stroke (Larson et al., [43]), multiple sclerosis (Aupperle et al., [1]; Beatty, [5]), schizophrenia (Wilk et al., [68]), and traumatic brain injury (TBI; McKay et al., [50]), among others. Consistent with broad neuropsychological literature, individuals with clinical conditions invariably perform worse on the RBANS subtests compared with the RBANS normative sample. Further indicative of the integration of this measure into routine neuropsychological practice, the RBANS served as a "gold standard" in a research study that evaluated the negative predictive power and positive predictive power of novel, brief, computerized neuropsychological assessment (Woodhouse et al., [69]).

The reliability of RBANS index scores has been evaluated in numerous studies. While reliability of the Total Score is generally high (Spearman Brown reliability coefficients range from .86 to .94), the reliability of the individual indexes was significantly lower (Spearman Brown reliability coefficients range from .55 to .78; Hobart, Goldberg, Bartko, & Gold, [29]; Randolph, [58]). McKay et al. ([50]) investigated the internal consistency of the RBANS indexes in a TBI sample and reported a wide range of alpha coefficients. While the Total Score ($\alpha = .83$), Delayed Memory ($\alpha = .77$), Visuospatial/Constructional ($\alpha = .76$), and Immediate Memory ($\alpha = .75$) indexes exhibited good internal consistency, the remaining RBANS indexes had unacceptable internal consistency (Attention, $\alpha = .16$; Language, $\alpha = .33$). Ultimately, despite regular use of the RBANS in clinical settings, low reliability of select indexes raises some question of whether select indexes (e.g., Attention and Language) evaluate a single latent construct.

Given that select RBANS indexes have lower-than-ideal reliability coefficients, it is not surprising that the RBANS factor structure is inconsistent with the index structure. Numerous researchers have investigated the internal structure of the RBANS, and these results are important to consider when evaluating the validity of test score interpretations. A clearly defined factor structure informs the fidelity of the construct-scoring structure and directly affects clinical decision making (King, Bailie, Kinney, & Nitch, [39]; Messick, [51]). While exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were not reported in the manual (Randolph, [58], [59]), to date, six studies investigating the factor structure of the RBANS have reported slightly inconsistent results (Carlozzi et al., [10]; Duff et al., [19]; Garcia, Leahy, Corradi, & Forchetti, [24]; King et al., [39]; Schmitt et al., [18]; Wilde, [67]).

Three CFA studies investigated the RBANS factor structure (Carlozzi et al., [10]; Duff et al., [19]; King et al., [39]). Each study evaluated whether the underlying factor structure of the RBANS was consistent with the RBANS index structure. Both a five-factor structure to mirror the index organization and a single-factor structure to replicate the overall score were investigated (Carlozzi et al., [10]; Duff et al., [19]; King et al., [39]). Across diverse samples, including community-dwelling older adults (Duff et al., [19]), veterans referred to a memory disorder clinic (Carlozzi et al., [10]), and patients with psychiatric disorders (King et al., [39]), CFA results did not support a five- or one-factor structure. Notably, across studies, immediate and delayed memory indexes were highly correlated, which contributed to a misfit between the underlying structure and expectation. This finding is not surprising given that numerous factor-analytic studies investigating memory have shown a single memory dimension that encompasses both immediate and delayed memory (Delis, Jacobson, Bondi, Hamilton, & Salmon, [16]; Dowling, Hermann, La Rue, & Sager, [17]; Hoelzle, Nelson, & Smith, [33]).

It is noteworthy that some researchers have expressed concern that CFA might not be an ideal method to evaluate validity evidence based on internal structure (Lee & Ashton, [44]). It has been observed that traditional fit indexes (e.g., χ^2 test) reject models that are only trivially misspecified when the sample size is large (Bentler & Bonett, [7]). Additionally, CFA may lack sensitivity for relationships between variables that may be highly discreet or complex because they must be specified a priori by the researcher (Hoelzle & Meyer, [32]). As an illustration, omnibus Big Five personality inventories have not been replicated when evaluated with CFA models (Church & Burke, [13]; Gignac, Bates, & Jang, [25]; McCrae, Zonderman, Costa, Bond, & Paunonen, [49]), despite the influence of factor-analytic methods on the development of the Big Five model of personality. RBANS CFA studies should

be interpreted with this in mind. In other words, the failure of CFA methods to support specified models does not necessarily mean the battery is invalid per se; rather, it raises questions about the relationship between subtests and the composition of indexes. This conclusion suggests alternative methods should be considered.

In contrast to theory-driven CFA, EFA is a data-driven method of variable reduction where multiple variables (e.g., subtests) are organized into factors or components that reflect relationships (e.g., cognitive constructs) between the variables (Goldberg & Velicer, [26]). EFA methods have been utilized to investigate the RBANS factor structure six times (see Table 1; Carlozzi et al., [10]; Duff et al., [19]; Garcia et al., [24]; King et al., [39]; Schmitt et al., [18]; Wilde, [67]). Close inspection of the pattern of factor loadings previously published reveals similarities and discrepancies across studies. Importantly, actual values of factor loadings vary depending on the methodology used, so specific loadings cannot be equated across all samples (Tabachnick & Fidell, [62]). Nevertheless, across studies, it appears that the key primary loadings on one factor typically reflect verbal memory functioning (List Learning, Story Memory, List Recall, List Recognition, Story Recall). The Figure Recall subtest loadings vary across studies between a first factor that is primarily a memory factor and a second factor typically reflecting visuospatial abilities or attention. The greatest discrepancy across studies appears to be how processing-speed, language, and attention tasks are associated with factors. This discrepancy is potentially troubling as these constructs are crucial to consider during the differential diagnosis process.

Table 1. Repeatable Battery for the Assessment of Neuropsychological Status factor-analytic studies overview.

Study	Sample	Mean age (SD)	Method	Rotation (subtests analyzed)	Extraction criteria	Latent constructs (% variance explained)
Wilde (2006)	210 patients with CVA	61.91 (13.97)	PCA	Varimax (12)	EV > 1 Scree plot	1. Language/verbal memory (37%)
2. Visual/visual memory (24%)						
Duff et al. (2006)	824 Normal Aging Adults	73.4 (5.8)	CFA ML EFA	Varimax Promax (9)	EV > 1 Scree plot	1. Verbal memory
2. Visual processing (60% combined)						
Garcia et al. (2008)	351 Memory Clinic patients	77.9 (7.5)	PCA	Direct oblimin (12)	EV > 1 Scree plot	1. Memory (39.5%)
2. Visuomotor processing (13.51%)						
3. Verbal processing (8.42%)						
Carlozzi et al. (2008)	175 Memory Clinic patients	74.1 (8.0)	CFA ML EFA	Varimax (12)	Chi-square test Variance explained	1. Memory, visual motor, verbal fluency (89.4%)
2. Visuospatial and attention (10.6%)						
Schmitt et al. (2010)	636 Memory Clinic patients	76.61 (7.29)	PCA PAFA	Varimax Promax (12)	EV > 1 Scree plot	1. Memory and learning
2. Visuospatial and attention (54.4% combined)						
King et al. (2012)	167 patients with SCZ	42.76 (9.73)	CFA PAFA PCA	Promax (12)	EV > 1 Scree plot SE of scree Horn's PA MAP	1. Memory (13.9%)
2. Speed of processing (8.2%)						

5 Note. CVA = cerebral vascular accident; SCZ = schizophrenia; PCA = principal components analysis; ML = maximum likelihood; EFA = exploratory factor analysis; EV = eigenvalue; CFA = confirmatory factor analysis; PAFA = principal axis factor analysis; SE of scree = standard error of the scree plot; PA = parallel analysis; MAP = minimum average partial.

Previous RBANS factor-analytic studies (Duff et al., [19]; Garcia et al., [24]; King et al., [39]; Schmitt et al., [18]; Wilde, [67]) have suggested that sample characteristics (e.g., normal cognitive functioning, memory impairment, psychiatric diagnoses) could meaningfully impact what cognitive constructs may emerge in EFA, a commonly held belief put forth by Delis, Jacobson, Bondi, Hamilton, and Salmon ([16]). For example, Garcia et al. ([24]) reported a three-factor EFA solution from a mixed clinical sample of outpatients with memory disorders and suggested this solution differed from others because of unique sample characteristics. Carlozzi et al. ([10]), Duff et al. ([19]), King et al. ([39]), and Wilde ([67]) reported two-factor solutions that differed slightly but also offered similar sample-based explanations for why discrepancies were observed in solutions. Simply stated, authors of previous RBANS factor-analytic studies proposed that solutions varied to a minor, but meaningful, degree because of underlying sample characteristics.

The issue of whether analyzing different samples results in different factor solutions has been thoroughly explored by researchers interested in measures that quantify mood and personality features. For example, O'Connor ([55]) investigated the factor structure of 37 different personality and psychopathology measures in clinical and nonclinical samples using empirically supported factor-analytic methods. O'Connor ([55]) conclusively identified that factor structures generally replicated across diverse samples. Hoelzle and Meyer ([31]) also found an invariant factor structure underlying the Personality Assessment Inventory (Morey, [52]) in different clinical and nonclinical samples after applying the same methods as O'Connor ([55]). Therefore, while researchers purport that different samples often yield different factor structures, it appears that this variability may actually reflect methodological decisions made by researchers and not solely underlying sample characteristics. The following section will briefly describe key EFA methodological issues that may be contributing to inconsistent factor solutions across different samples.

Exploratory factor-analytic methodology

It is often overlooked that there are multiple ways to conduct EFA. Factor analysis (FA) and principal components analysis (PCA) are both data-driven approaches to identify underlying dimensions, but they differ in theory. The primary mathematical difference is what value is placed on the main diagonal of the correlation matrix (Goldberg & Velicer, [26]; Tabachnick & Fidell, [62]). In FA, the covariance between variables is analyzed so error and unique variance are excluded from *factors*; thus, values in the correlation matrix diagonal are communalities of the shared variance between variables (e.g., values between 0 and 1). In PCA, 1 is in the diagonal of the correlation matrix and all variance, including error and unique variance, is disseminated to the *components*. Because error and unique variance are omitted in FA, the observed variables and observed correlation matrix are not fully reproduced and the factors are approximates. It is suggested that the difference between the two methods does not meaningfully impact results (Goldberg & Velicer, [26]; Hoelzle & Meyer, [32]). Consistent with this position, two RBANS factor-analytic studies showed that when both FA and PCA were conducted, similar results were obtained (Carlozzi et al., [10]; King et al., [39]). This finding has also been observed when researchers investigated self-report and performance-based measures (e.g., see Nelson, Sweet, Berry, Bryant, & Grancher, [53]). Therefore, whether FA or PCA was conducted should not result in meaningful differences between RBANS factor solutions.

A key methodological decision when conducting FA is determining how many factors to retain. Four previous factor-analytic RBANS studies (Duff et al., [19]; Garcia et al., [24]; Schmitt et al., [18]; Wilde, [67]) utilized the two most common methods to determine factor retention: *Kaiser's criterion* (Kaiser, [37]) and visual examination of a scree plot (Cattell, [12]). The methods used by researchers to investigate the RBANS factor structure are somewhat inconsistent with best-practice guidelines (Fava & Velicer, [22], [23]; Goldberg & Velicer, [26]; Hoelzle & Meyer, [32]; Hubbard & Allen, [36]; Zwick & Velicer, [70], [71]). In brief, empirical research suggests that multiple methods of factor extraction should be utilized to identify a reliable factor solution including: interpretation of the scree plot, Horn's ([35]) parallel analysis (PA), and the minimum average partial (MAP) procedure (Velicer, [63]). King et al. ([39]) utilized these retention methods; however, the results diverged. Kaiser's criterion indicated that two factors should be retained, and all other methods suggested a one-factor solution. Despite strong evidence from multiple empirically supported factor-retention procedures that one factor should be retained, King et al. ([39]) selected a final solution that was supported only by Kaiser's criterion. It is significant that researchers have not uniformly utilized empirically supported guidelines to determine how many factors to retain. Research has suggested that neglect of empirical guidelines for factor retention can result in inconsistent findings across studies (Hoelzle & Meyer, [20]; O'Connor, [55]). Thus, it is possible that if empirically supported procedures were uniformly implemented, subtests would load more consistently on factors (or a single factor) and a truly invariant RBANS factor structure would be more likely to emerge.

Determining how extracted factors will be rotated prior to interpretation is also an important methodological decision, and recommendations clearly indicate that when factors (e.g., distinct cognitive constructs) are known to be correlated, oblique rotation (e.g., direct oblimin) should be selected (Carroll, [11]; Deary, [15]; Hoelzle et al., [33]; Tabachnick & Fidell, [32]). Interestingly, only one RBANS factor-analytic study utilized an oblique method of rotation (Garcia et al., [24]). Four of the prior studies (Carlozzi et al., [10]; Duff et al., [19]; Schmitt et al., [18]; Wilde, [67]) utilized orthogonal rotation (e.g., Varimax), which assumes that factors are uncorrelated. This decision is questionable because by nature, cognitive constructs are correlated with each other (e.g., attention is meaningfully related to memory functioning). A third rotation, Promax, which involves aspects of oblique and orthogonal rotation, was utilized in three studies (Duff et al., [19]; King et al., [39]; Schmitt et al., [18]). It is unclear to what degree differences in factor rotation have possibly contributed to differences between solutions. In fact, oblique and orthogonal rotations will result in similar solutions if identified factors are truly not correlated with one another. Nevertheless, there is a strong theoretical rationale for using oblique rotation given the well-documented relationships between cognitive abilities.

Current study

Based on a review of literature, it is clear that slightly discrepant RBANS factor structures have been reported. A common factor emerges across studies that reflects the latent construct of memory, but while the second factor primarily emphasizes visuospatial functioning, some subtests meaningfully loaded on different dimensions across studies. Although researchers suggest discrepant findings are related to sample-based issues, there is a body of literature that suggests methodological issues, specifically factor-retention decisions, primarily contribute to these differences. The overarching goal of the present study was to quantify the similarity of solutions across multiple samples after systematically analyzing RBANS data sets using empirically supported methods. The congruence of

factor solutions will be investigated by conducting orthogonal vector matrix comparisons (Barrett, [2]). The outcome of this study will synthesize previous factor-analytic investigation of the RBANS and will evaluate the invariance of solutions. Due to the expansive use of this neuropsychological instrument, a definitive conceptualization of the RBANS may have significant clinical implications in that it could definitively clarify the relationships between subtests and indexes. In addition, this research provides an example of how empirically supported extraction can be applied to multiple samples and factor structures can be systematically compared.

Method

Samples and procedures

Institutional review board approval was obtained for this archival study from both the Gundersen Health System and Marquette University. A new clinical sample was obtained from the Gundersen Health System Memory Clinic (La Crosse, WI). This sample included 393 patients who were evaluated from January 1, 2009, to June 1, 2013, by a multidisciplinary team. Participants with significant cognitive impairment (e.g., Mayo Short Test of Mental Status [Kokmen, Naessens, & Offord, 1987] score <14 or severe intellectual disability, $n = 48$) were administered an abbreviated neuropsychological battery that did not include the RBANS and were therefore excluded from this study. Patients included in this sample ($n = 345$) ranged in age from 44 to 96 years old ($M_{\text{age}} = 75.29$ years, $SD = 8.68$). Fifty-three percent ($n = 186$) of the sample was female. Estimates of premorbid intellectual functioning indicated this sample was within the average range (Wechsler Test of Adult Reading, Psychological Corporation, [57], $n = 130$, $M = 95.43$, $SD = 15.66$; ACS Test of Premorbid Functioning, Wechsler, [65], $n = 217$, $M = 94.08$, $SD = 11.23$). The majority of this sample completed high school ($M = 12.66$, $SD = 3.10$). This sample was diverse diagnostically, though the majority of patients received a diagnosis of dementia (Alzheimer disease, 24.9%; dementia-not otherwise specified [NOS], 18.8%; cognitive disorder-NOS, 13.2%; vascular dementia, 10.9%; frontotemporal dementia, 7.6%; mild cognitive impairment, 7.1%; normal/no impairment, 5.1%; mixed dementia, 5.1%; Lewy body dementia, 2.3%; Parkinson dementia, 1.5%; pervasive developmental disability, 1.5%; attention-deficit hyperactivity disorder, 0.8%; Wernicke-Korsakoff, 0.3%). Racial and ethnic identity was not reliably available for this sample in electronic medical records, though the sample was predominantly Caucasian and not of Hispanic origin. This sample was included to avoid the occurrence of sample-specific floor or ceiling effects because restriction in range may impact the possible strength of factor loadings and correlations (Fabrigar, Wegener, MacCallum, & Strahan, [21]; Larrabee, [42]).

Previously published RBANS data were also reanalyzed. Carlozzi et al. ([10]) published their RBANS correlation matrix, so it was possible to include that data in analyses. Wilde ([67]) had previously supplied a correlation matrix for a prior research project (Hoelzle, [30]), so that sample is also included. Through personal communication, RBANS correlation matrixes were requested from each of the corresponding authors of the remaining four RBANS factor-analytic studies (Duff et al., [19]; Garcia et al., [24]; King et al., [39]; Schmitt et al., [18]). One RBANS correlation matrix previously analyzed was obtained (Duff et al., [19]), and that primary author also provided an additional correlation matrix from a sufficiently large sample (Duff et al., [18]). Basic information regarding characteristics of the samples in Carlozzi et al. ([10]), Duff et al. ([19]), and Wilde ([67]) are presented in Table 1. Of note, the correlation matrix provided by Duff et al. (2006) had a slightly different sample size ($N = 796$) than was

reported in the 2006 publication ($N = 824$). Similarly, Duff provided a correlation matrix consisting of data obtained from a sample consisting of individuals who were cognitively intact or had mild cognitive impairment ($N = 173$), which was larger than what was reported in the corresponding publication ($N = 143$; Duff et al., [18]). Therefore, the sample characteristics for these two samples are likely to minimally diverge from what was published by Duff and colleagues ([19]).

Procedures: Statistical analysis

PCA was conducted to evaluate the underlying dimensional structure of each sample. Although this method technically extracts *components*, the term *factor* will be used interchangeably because it is common in the literature. Prior to conducting analyses, the Kaiser-Meyer-Olkin (KMO) statistical index was reviewed to evaluate whether there was problematic collinearity between variables (e.g., $KMO < .70$; Kaiser, [38]). All samples were appropriate for PCA (Carlozzi et al., [10], $KMO = .91$; Duff et al., [19], $KMO = .88$; Duff et al., [18], $KMO = .83$; Wilde, [19], $KMO = .87$; Gundersen, $KMO = .87$).

The key methodological decision in EFA is determining how many factors to extract from the observed correlation matrix. Although extracting too many factors may result in a solution that more closely recreates the original correlation matrix, it increases the odds that meaningful factors will split and result in unreliable components (Fava & Velicer, [22]). If a parsimonious solution is sought, the investigator may risk extracting too few factors by combining distinct components and oversimplifying the solution (Fava & Velicer, [22]). Employing multiple empirically supported extraction techniques improves the likelihood that a reliable solution will emerge across diverse samples (Fabrigar et al., [21]; Hoelzle & Meyer, [31]; O'Connor, [55]).

A simple procedure often used to guide retention decisions is Kaiser's criterion, which states that all components with eigenvalues greater than 1 should be retained (Kaiser, [37]). The problem with this approach is that the number of eigenvalues greater than 1 is typically equal to a value that is one fifth to one third of the total number of variables analyzed regardless of the actual underlying structure of data (Zwick & Velicer, [70]). Empirical research has demonstrated that Kaiser's criterion regularly results in overextraction and inconsistent component solutions (Fabrigar et al., [21]; Hubbard & Allen, [36]; Preacher & MacCallum, [56]; Zwick & Velicer, [70]).

Visual examination of the scree plot or eigenvalue plot is another frequently utilized technique to guide factor extraction decisions (Cattell, [12]). Although this approach works well when there are unique factors that account for significant amounts of variance, the technique tends to be highly subjective when factor differentiation is weak (Linn, [46]; Zwick & Velicer, [70]). Therefore, researchers suggest that alternative factor extraction or retention guidelines should be utilized as well (Goldberg & Velicer, [26]).

PA also examines eigenvalues but is considered a more reliable technique because sampling error is considered (Horn, [35]). PA involves generating correlation matrixes from random data that include the same number of variables and subjects as the actual correlation matrix. The eigenvalues from the randomly generated data are then compared to the actual eigenvalues, and only factors with eigenvalues greater than those from the random data are retained. Simulated empirical investigations have shown that PA is one of the most accurate methods in determining the dimensions present in PCA (Crawford et al., [14]; Zwick & Velicer, [71]).

The MAP procedure is an alternative extraction technique that was initially designed for PCA (Velicer, [63]). The MAP procedure sequentially removes each component from the original correlation matrix and then creates a partial correlation matrix. As each component is removed, the average of the squared partial correlations is computed. As long as each component contains common variance, the average of the squared partial correlations decreases. This value increases when an extracted component consists of unique variance and at that point suggests overextraction. In other words, the suggested number of components to retain is determined at the point at which the average squared partial correlation is smallest. Empirical research has determined that the MAP procedure is the most reliable extraction technique (Zwick & Velicer, [70], [71]).

In summary, researchers have followed a number of different procedures to determine how many factors should be extracted in PCA. Unfortunately, the methods most often utilized—Kaiser's criterion and the interpretation of scree plots—are most likely to result in nonreplicating solutions. Factor-retention decisions in the present study were based on PA and MAP procedure results.

After determining how many factors will be extracted, the next step is to rotate the matrix of loadings to aid interpretability (Goldberg & Velicer, [26]). As noted previously, the decision was made to implement oblique (direct oblimin) rotation for theoretical reasons. Finally, factor solutions were carefully reviewed to determine what latent constructs have been identified. Interpretatively, items with strong loadings will reflect the cognitive construct, whereas those variables with loadings near 0 will indicate the absence of a construct.

Utilization of empirically validated extraction methods does not quantify the similarity of solutions obtained from different samples. Orthogonal vector matrix comparisons (Barrett, [2]) evaluate the similarity of solutions across samples beyond a visual examination of loadings. Orthogonal vector matrix comparison methods rotate one sample structure to align it with a solution from another sample (Barrett, [2]; Barrett, Petrides, Eysenck, & Eysenck, [4]). Rotation occurs to maximally align the solutions without distorting the original component solutions when a sample solution is compared to a target solution (Barrett, [2]; Barrett et al., [4]; Hoelzle & Meyer, [31]; Hopwood & Donnellan, [34]). This comparison method results in congruence coefficients that indicate how well factors match one another.

Comparison of single-component structures requires a different statistical process than comparison of multidimensional component structures. Tucker's congruence coefficient accounts for both the pattern and magnitude of loadings to determine if a single-factor solution is replicated across samples (Korth & Tucker, [41]; Levine, [45]). Additionally, single-component structures can be compared using Pearson's r when a solution has few small loadings (<.40) to, again, compare pattern and magnitude of loadings. Multiple small loadings within a factor will generate a large r value masking the impact of more significant loadings, so caution is warranted when utilizing Pearson's r to quantify the similarity of single-factor solutions (Lorenzo-Seva & ten Berge, [47]).

Results

PCA was conducted separately for each sample. Four factor-retention guidelines were applied, though PA and the MAP procedure were most strongly considered when determining how many factors to extract from each sample. The respective number of components suggested by each procedure are

presented in Table 2. Not surprisingly, given limitations previously discussed, Kaiser's criterion and visual examination of the scree plot resulted in discrepant recommendations regarding how many factors to retain across samples.

Table 2. Principal components analysis extraction criteria results summary.

	Carlozzi et al. (2008)	Duff et al. (2010)	Duff et al. (2006)	Wilde (2006)	Gundersen
EV > 1	2	4	2	2	3
Scree plot	1	2	3	2	1
PA	1	1	2	2	2
MAP	1–2	1	1–2	2	1–2

6 *Note.* EV = eigenvalue; PA = parallel analysis; MAP = minimum average partial. MAP ranges reflect minor differences between MAP procedures not exceeding .04.

PA was conducted using O'Connor's ([54]) syntax, and the results are presented in Table 3. In these analyses, 500 random data sets were generated with the same number of test variables (e.g., 12 RBANS subtests) and matched sample size. The actual eigenvalues were compared to the 95th percentile of randomly generated eigenvalues (as opposed to mean eigenvalue) to decrease the risk for overextraction in situations when sample sizes are small and expected factor loadings are low (Zwick & Velicer, [71]). In the present study, PA indicated retaining one factor in two samples (Carlozzi et al., [10]; Duff et al., [18]) and two factors in the other three samples (Duff et al., [19]; Wilde, [67]; Gundersen). Interpretatively, retention recommendations would not have changed if actual eigenvalues were compared to the mean PA eigenvalues as opposed to the 95th percentile of randomly generated eigenvalues.

Table 3. Repeatable Battery for the Assessment of Neuropsychological Status actual and random eigenvalues from Horn's parallel analysis.

	Carlozzi et al. (2008)		Duff et al. (2010)		Duff et al. (2006)		Wilde (2006)		Gundersen	
	EV	PA EV	EV	PA EV	EV	PA EV	EV	PA EV	EV	PA EV
1	6.51	1.56	4.63	1.57	5.00	1.25	5.33	1.51	5.27	1.39
2	1.02	1.41	1.20	1.42	1.35	1.19	1.98	1.37	1.38	1.28
3	0.89	1.30	1.07	1.30	0.98	1.14	0.90	1.27	1.06	1.21
4	0.71	1.22	1.02	1.21	0.82	1.10	0.69	1.20	.82	1.15
5	0.58	1.13	0.90	1.13	0.74	1.07	0.65	1.12	0.67	1.11
6	0.47	1.06	0.76	1.07	0.59	1.03	0.50	1.07	0.57	1.05
7	0.41	1.00	0.66	1.00	0.56	1.00	0.46	1.00	0.55	1.01
8	0.38	0.94	0.49	0.94	0.52	0.98	0.40	0.94	0.46	0.96
9	0.35	0.88	0.44	0.88	0.49	0.95	0.35	0.89	0.39	0.92
10	0.27	0.81	0.38	0.81	0.45	0.91	0.30	0.83	0.37	0.87
11	0.25	0.75	0.27	0.75	0.32	0.88	0.24	0.78	0.31	0.82
12	0.17	0.68	0.18	0.68	0.19	0.84	0.21	0.71	0.16	0.77
Retained	1		1		2		2		2	

7 *Note.* Real EV = actual data eigenvalue; PA EV = 95th percentile eigenvalue of randomly generated data. Bold and italic values indicate the number of components recommended for retention.

Velicer's (1976) MAP procedure was conducted using O'Connor's ([54]) syntax. Results of the MAP procedure are visually presented in Figure 1. MAP indicated retaining one factor in Duff et al. ([18]) and two factors in Wilde ([19]). The other three samples (Carlozzi et al., [10]; Duff et al., [19]; Gundersen) had average partial correlations that were extremely close (e.g., $<.04$), which suggests that one- and two-factor solutions should be further considered.

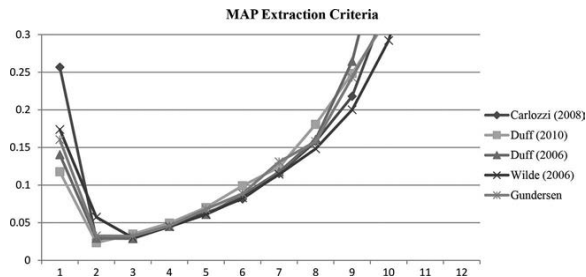


Figure 1. Velicer's minimum partial average procedure indicating number of components to be retained for each Repeatable Battery for the Assessment of Neuropsychological Status sample.

PCA was conducted specifying a two-component solution for each sample, and factor loadings for each sample are presented in Table 4. An examination of solutions revealed similarity across samples. The first dimension strongly suggested a memory construct (List Recall, Story Recall, List Learning, List Recognition, Story Memory, Semantic Fluency, and Figure Recall) and accounted for at least 38.54% of the total score variance. The Semantic Fluency subtest, a verbal fluency task that involves rapidly recalling information from specific categories, also reliably loaded on the first factor. It is notable that the Figure Recall subtest displayed meaningful cross-loading in two samples, and in the Wilde ([67]) sample, the subtest was strongly associated with a nonmemory dimension. Nevertheless, most reliably, Figure Recall was associated with Factor 1.

Table 4. Two-component Repeatable Battery for the Assessment of Neuropsychological Status oblique rotated pattern matrixes.

	Carlozzi et al. (2008)		Duff et al. (2010)		Duff et al. (2006)		Wilde (2006)		Gundersen	
	1	2	1	2	1	2	1	2	1	2
List Learning	.75	.14	.85	.06	.75	.09	.86	-.03	.76	.12
Story Memory	.77	.11	.72	.06	.75	.11	.78	-.09	.75	.12
Figure Copy	-.07	.85	-.22	.81	-.05	.76	-.15	.96	-.12	.80
Line Orientation	-.04	.89	.07	.48	-.16	.83	-.04	.79	-.01	.81
Picture Naming	.20	.56	.23	.25	.03	.62	.69	-.03	.24	.35
Semantic Fluency	.65	.14	.60	.11	.40	.23	.70	.08	.62	.18
Digit Span	.06	.51	.12	.38	.09	.37	.50	-.02	.23	.14
Coding	.48	.48	.59	.27	.24	.59	.31	.66	.22	.70
List Recall	.94	-.22	.87	-.18	.89	-.10	.74	.08	.89	-.17
List Recognition	.66	.16	.83	-.28	.80	-.12	.80	-.01	.78	-.08
Story Recall	.89	-.02	.81	-.20	.82	.08	.77	.09	.91	-.08
Figure Recall	.45	.42	.60	.15	.36	.46	.10	.79	.74	.01
Eigenvalue	6.51	1.02	4.63	1.20	5.00	1.35	5.33	1.98	5.27	1.38
r between factors	.61		.53		.48		.36		.44	
% variance explained	54.27	8.49	38.54	10.02	41.70	11.25	44.41	16.49	43.92	11.50

Factor 2 appeared to reflect a visuospatial construct, and the Figure Copy, Line Orientation, and Coding subtests reliably had high loadings. The Coding subtest had meaningful factor loadings on both dimensions. These cross-loadings might be attributed to the attentional component required in Coding that is conceptually similar to the attention requirements in list- and story-learning tasks. Additionally, Coding and Semantic Fluency possess a mutual speed component. The second visuospatial factor accounted for 9% to 17% of the total score variance, which is meaningfully less than the first factor. The two remaining subtests, Picture Naming and Digit Span, did not consistently load on either factor.

Additional analyses were conducted to derive and evaluate one-factor RBANS solutions because PA and the MAP procedure provided some support for retaining only one factor in several samples. The majority of RBANS subtests meaningfully loaded onto the factor (see Table 5). Subtests with the strongest loadings were generally memory tasks indicating the primary presence of the cognitive construct of memory, though language, processing speed, and perceptual organization are also meaningfully emphasized. Digit Span had relatively low loadings on the one-factor solution (pattern matrix loadings $<.40$) in three samples. The amount of variance explained in the single-factor solution mirrors the amount of variance explained by the memory factor in the two-factor solution (Factor 1 variance $\geq 38.54\%$).

Table 5. Repeatable Battery for the Assessment of Neuropsychological Status single-component solution.

	Carlozzi et al. (2008)	Duff et al. (2010)	Duff et al. (2006)	Wilde (2006)	Gundersen	Averaged loadings
List Learning	.82	.86	.76	.80	.81	.81
Story Memory	.82	.74	.78	.68	.80	.76
Figure Copy	.65	.07	.55	.50	.41	.44
Line Orientation	.71	.23	.50	.51	.52	.49
Picture Naming	.66	.32	.51	.63	.46	.52
Semantic Fluency	.72	.64	.55	.71	.72	.67
Digit Span	.49	.25	.36	.46	.31	.37
Coding	.86	.68	.67	.72	.67	.72
List Recall	.70	.80	.74	.75	.75	.75
List Recognition	.75	.73	.64	.75	.70	.71
Story Recall	.82	.80	.82	.79	.83	.81
Figure Recall	.78	.65	.70	.61	.72	.69
Eigenvalue	6.51	4.63	5.00	5.33	5.27	5.35
Percent of variance explained	54.27	38.54	41.70	44.41	43.92	44.57

Next, quantitative methods of factor comparison were conducted using Orthosim 2.1 (Barrett, [3]) and are displayed in Table 6. Recommendations for interpreting congruence coefficients vary. Barrett, Petrides, Eysenck, and Eysenck ([4]) suggested coefficients in the range of .80 to .95 demonstrate good similarity of factors and coefficients at .98 and greater indicate an identical factor between samples. More delineated interpretive guidelines have been put forth as well: excellent = .98 to 1.00, good = .92 to .98, borderline = .82 to .92, poor = .68 to .82, and terrible < .68 (MacCallum, Widaman, Zhang, & Hong, [48]). Resulting congruence coefficients vary slightly depending on which sample is the target sample, so all congruency coefficients are reported in Table 6. Overall, there was strong support for the similarity of a two-component solution with all coefficients except 1 meeting Barrett et al.'s ([4]) guidelines for factor replication. In addition, when considering the delineated guidelines, nearly 75% of the coefficients (33 out of 40) met MacCallum et al.'s ([48]) benchmarks for "good" or "excellent" replication. This evidence is sufficient to support the position that there is an invariant two-factor RBANS structure that emerges across samples. Despite there being subtle differences in factor loadings across subtests (e.g., Picture Naming and Digit Span), it did not negatively impact the congruency of solutions.

Table 6. Two-component vector matrix comparisons with 12 Repeatable Battery for the Assessment of Neuropsychological Status subtests.

	Carlozzi et al. (2008)		Duff et al. (2010)		Duff et al. (2006)		Wilde (2006)		Gundersen	
	1	2	1	2	1	2	1	2	1	2
Carlozzi et al. (2008)	—	—	.96	.94	.95	.93	.95	.88	.98	.95
Duff et al. (2010)	.98	.94	—	—	.97	.97	.93	.87	.98	.94
Duff et al. (2006)	.99	.96	.97	.95	—	—	.95	.90	.99	.96
Wilde (2006)	.94	.91	.95	.84	.93	.92	—	—	.95	.88
Gundersen	.98	.96	.98	.92	.98	.96	.95	.90	—	—

For the sake of completeness, the similarity of one-component solutions across samples was also investigated (see Table 7). Tucker's congruence coefficients (Levine, [45]; Lorenzo-Seva & ten Berge, [47]) suggest excellent congruence of one-factor solutions (e.g., interpretative benchmarks "similar" = .85–.94, "identical" = .95–1.00). Additionally, Pearson *r* correlations were calculated to quantify congruency of single-factor solutions (Levine, [45]), and all solutions were highly correlated. Both procedures indicated a single-factor RBANS dimension is invariant across samples.

Table 7. Single-component solution comparisons with 12 Repeatable Battery for the Assessment of Neuropsychological Status subtests.

	Carlozzi et al. (2008)	Duff et al. (2010)	Duff et al. (2006)	Wilde (2006)	Gundersen
Carlozzi et al. (2008)	—	.72**	.84**	.74**	.85**
Duff et al. (2010)	.94	—	.83**	.92**	.93**
Duff et al. (2006)	.99	.96	—	.80**	.91**
Wilde (2006)	.99	.96	.99	—	.88**
Gundersen	.99	.98	.99	.99	—

8 Note. Tucker's congruence coefficients are located below the diagonal and Pearson's *r* values are above the diagonal. ** $p < .01$.

Discussion

Neuropsychological test validation is a continual process that requires examination of a measure utilizing different clinical and nonclinical samples. The factor structure of a measure is directly related to the validity of test score interpretation and thus is an important area of focus for researchers and clinicians. The present study investigated the factor-analytic structure of the RBANS, a widely used neuropsychological measure (e.g., see Randolph, [58], [59]). To date, six studies have evaluated the RBANS factor structure and revealed slightly different solutions (Carlozzi et al., [10]; Duff et al., [19]; Garcia et al., [24]; King et al., [39]; Schmitt et al., [18]; Wilde, [67]). This extensive body of literature clearly indicates that the RBANS factor structure is inconsistent with the theoretically developed RBANS five-index and single neuropsychological score structure. However, direct comparison of previous factor-analytic solutions is somewhat confounded because different methods were utilized. While many researchers have suggested that divergent factor-analytic solutions are associated with differences in sample characteristics, this research was conducted to determine (a) the degree to which methodological decisions impact findings by uniformly applying EFA methods, and (b) the quantitative similarity of solutions across diverse samples.

Consistent with expectations, Kaiser's criterion and visual examination of the scree plot provided variable factor-retention recommendations both across and within samples (see Table 2). Horn's ([35]) and Velicer's ([63]) MAP procedures reliably indicated retention of one or two factors. Hence, both two- and one-factor solutions were explored to alleviate risk for overextraction or underextraction.

With respect to a two-factor solution, this data-driven investigation revealed a strong first component of memory (List Recall, Story Recall, List Learning, List Recognition, Story Memory, Semantic Fluency, and Figure Recall) and a second visuospatial component (Figure Copy, Line Orientation, and Coding). At face value, these results are consistent with those of previous studies (Carlozzi et al., [10]; Duff et al., [19]; King et al., [39]; Schmitt et al., [18]; Wilde, [67]), but the current analyses permitted a more

refined understanding of the relationship among subtests and underlying dimensions. While previous researchers have attributed minor differences to unique sample characteristics, this research evaluated the similarity of solutions after uniformly applying quantitative methods. The congruency of two-factor solutions suggests that differences between samples are inconsequential and a robust, invariant two-dimensional structure underlies the RBANS, which is generally consistent with previous findings.

Not surprising given previous findings (Carlozzi et al., [10]; King et al., [39]), the Picture Naming and Digit Span subtests did not consistently load on either identified factor (see Table 4); however, this finding minimally impacted the overall congruency of solutions because the corresponding loadings were not prominent in defining factors. Duff et al. ([19]) eliminated the Picture Naming subtest prior to analyses due to a ceiling effect and, additionally, removed Digit Span from analyses due to low correlations with other subtests. In contrast, Digit Span and Picture Naming appeared to reliably load onto the second factor in studies by Garcia et al. ([24]), Schmitt et al. ([18]), and Wilde ([67]). To further explore how these subtests impact the overall congruency of solutions, these subtests were excluded and analyses were repeated. Overall, congruency between solutions improved (average Factor 1 congruency coefficient = .98; average Factor 2 congruency coefficient = .95; all coefficients $\geq .90$), and this finding confirmed that these subtests contributed to lower-than-exceptional congruency of factor solutions derived from different samples.

The RBANS Picture Naming and Digit Span tasks are conceptually expressive vocabulary and attention tasks, respectively. The cognitive skills necessary to complete these tasks are not theoretically related to the primary constructs associated with the identified factors. Further, empirical investigation reveals that a minimum number of marker variables must be present for a related component to emerge in FA (Velicer & Fava, [64]). Consequently, there simply is not sufficient representation of confrontation naming and attentional skills to enable these subtests to load reliably onto a corresponding factor. Related to this issue, the Attention and Language indexes did not emerge in this study nor in previous factor-analytic studies (Carlozzi et al., [10]; Duff et al., [19]; Garcia et al., [24]; King et al., [39]; Schmitt et al., [18]; Wilde, [67]) due to typically low (or at best moderate) relationships between subtests that comprise these indexes. Previous RBANS literature has revealed poor internal consistency of the Attention and Language indexes (Beatty, [5]; Beatty et al., [6]; Larson et al., [43]; McKay et al., [50]), and the current research raises questions regarding interpretation of these theoretical indexes. There is limited support for the interpretation of the Attention and Language indexes based on the internal structure of the RBANS. Although it is not surprising that the Picture Naming and Digit Span subtests did not meaningfully load with identified factors, it does not mean that the subtests lack clinical utility per se. However, clinicians should recognize that interpretation of a single subtest to represent a cognitive construct is neither optimally reliable nor sensitive and may ultimately negatively impact the clinical utility of the measure.

A closer investigation of Picture Naming reveals that a ceiling effect is clearly present. The mean scores minimally ranged across samples ($M = 8.87-9.56$; $SD = 0.73-1.53$). This finding is meaningful because a restricted score range attenuates relationships between that task and other subtests (Fabrigar et al., [21]). In other words, a subtest with a skewed distribution of scores is limited in its ability to meaningfully correlate with other subtests (Goldberg & Velicer, [26]) and, ultimately, how it might load

on identified dimensions. Additionally, a ceiling effect suggests potentially limited clinical utility. Clinicians should recognize that the subtest may only effectively identify expressive language issues that are readily apparent to the examiner. In current form, it is questionable whether the subtest would detect mild language issues. A practical implication of this observation may involve revisions to the Picture Naming subtest, which certainly presents challenges due to the strong associations between this task, education, and culture. A novel approach to item selection, test length, and basal/ceiling rules could be determined by identifying the item parameters of a larger pool of items utilizing item response theory. Calamia, Markon, Denburg, and Tranel ([8]) illustrated the value of this approach by developing a robust short form of Judgment of Line Orientation that is sensitive to visuospatial difficulties and is meaningfully related to the full version of the original test. Although such methods would not increase the likelihood of Picture Naming loading on one of the two identified factors, it is quite possible that the method would result in increased sensitivity to language issues.

With respect to Digit Span, in two previous factor-analytic studies, the Digit Span subtest was not meaningfully related to identified factors (Carlozzi et al., [10]; King et al., [39]). Examination of the RBANS two-component solution consistency reveals low loadings for the Digit Span subtest and inconsistency in loading on either factor. As noted previously, the RBANS Digit Span task is conceptually an attention task. Anecdotally, in clinical settings, tasks assessing working memory, executive functioning, and attention are frequently administered in addition to the RBANS. A recommended revision to the RBANS is to include expansion of the Digit Span task to include backward and sequencing components (similar to the Wechsler Adult Intelligence Scale-Fourth Edition Digit Span subtest; Wechsler, [66]). To further develop a working-memory component, an additional working-memory task, such as mental arithmetic or letter-number sequencing, could also be added to the RBANS. More thorough assessment of working memory could improve clinical utility of the RBANS across diverse populations, as this construct is often impaired in psychiatric (e.g., anxiety and mood disorders) and neurologic conditions (e.g., attention-deficit hyperactivity disorder, dementias, mild TBI). Revisions to more thoroughly evaluate attention, divided attention, and working memory are likely to improve clinical utility and make the RBANS more useful across multiple populations while still keeping test administration relatively brief.

An invariant two-factor RBANS structure has implications for future research and clinical practice. Component scores could be developed using a normative sample and a unit-weighting scheme or exact factor score approach (Grice, [27], [28]). Each of these procedures could be explored to determine whether the empirically derived factor scores or theoretically developed index scores are more useful in detecting cognitive impairment or meaningful change from a baseline level of functioning. Empirically derived factor scores may have greater clinical utility because theoretically, they should have improved reliability and therefore be more sensitive to change. Duff et al. ([20]) recognized the likely presence of a two-factor solution and proposed a method to derive Verbal Index (List Learning, Story Memory, List Recall, List Recognition, and Story Recall) and Visual Index (Figure Copy, Figure Recall, Line Orientation, and Coding) scores. While the present study strongly supports consideration of the Memory and Visuospatial Indexes, it is noteworthy that Duff et al.'s ([20]) method assigns equal score weighting to subtests regardless of how they define a factor. A "true" factor score would reflect the weight of each subtest within the dimension and thus would more accurately reflect the component solution (Grice, [27], [28]). It would be worthwhile to consider how interpretation of the

subtests, Duff and colleagues' ([20]) indexes, or a purely, empirically derived factor score approach could predict outcomes in clinical samples with well-defined impairment affecting the respective constructs.

Admittedly, this research was conducted with the assumption that empirically supported retention procedures would likely recommend the retention of a single-factor solution based on findings from King et al. ([39]) because they incorporated PA and MAP procedures, and both procedures suggested the retention of a single factor. It was hypothesized that this systematic review and investigation would produce results that are more inconsistent with the literature than they are. Given support for the retention of just one factor in select samples (see Table 2), one-factor solutions were also considered. Examination of these solutions revealed that the majority of subtests loaded onto the single component with the exception of Digit Span. Additionally, quantitative analysis of factor congruency across samples revealed strong evidence for solution replication. However, when a one-factor solution was specified, the subtests that most strongly defined the dimension were tasks associated with memory functioning. It appears that the primary underlying cognitive construct of a single empirically derived factor score of the RBANS is not general neuropsychological status, but rather, predominantly memory functioning. An empirically derived factor score would be most sensitive to memory deficits as opposed to other cognitive issues. A single-factor structure, in comparison with the robust two-factor solution, compresses RBANS subtests that evaluate visuospatial functioning. This finding reveals a clear disadvantage for the one-factor solution. Moreover, one could argue that a two-factor solution is more informative and would have greater clinical utility. For example, it would be more likely to detect visuospatial deficits associated with a right parietal cerebral vascular accident.

This review also serves as an example of how researchers might evaluate the congruency of factor solutions across multiple samples. There are numerous examples in the literature where authors have identified different factor solutions across samples and raised questions regarding the validity of test score interpretations based on the relationship between aspects of the test and underlying factors (O'Connor, [55]). There is an emerging literature suggesting that it is essential to consider decisions regarding factor extraction when conducting EFA. As an exercise to demonstrate the potential drawback of overextraction, additional analyses were conducted to explore a three-factor RBANS solution. One previous study revealed a three-factor solution (Garcia et al., [24]), and two extraction criteria (e.g., Kaiser's criterion and visual examination of the scree plot) indicated the possibility of retaining three factors. A three-factor solution clearly resulted in nonreplicating solutions because the third factor was typically only defined by Digit Span, and the other RBANS subtests shifted between factors in an inconsistent manner. Orthogonal vector matrix comparisons indicated poor replication of a three-factor solution (mean Factor 1 congruency coefficient = .97; mean Factor 2 congruency coefficient = .91; mean Factor 3 congruency coefficient = .80). Additionally, approximately half of the congruency coefficients were in the borderline, or lower, range (26/60) and only a small number of coefficients were in the exceptional range (8/60). These findings clearly demonstrate the importance of utilizing empirically supported factor-retention strategies (e.g., PA and the MAP procedure) to identify an invariant factor structure.

Conclusions

Widespread agreement exists that a viable and defensible factor structure does not emerge from a single analysis. An optimal factor structure is one that is replicated across multiple diverse samples, with varying sample size and characteristics (Goldberg & Velicer, [26]). The present study has documented an invariant two-component RBANS solution through exploratory analysis and confirmed pattern replication by conducting orthogonal vector matrix comparisons. These factors primarily reflect memory and visuospatial cognitive constructs. Furthermore, the Picture Naming and Digit Span subtests were minimally correlated with other RBANS subtests, and as a result, they did not consistently load onto the replicated factors. Although not initially hypothesized, these results were generally consistent with previous studies and therefore provide further confirmation of a solution that is inconsistent with the theoretical RBANS index structure. This research quantifies the similarity of factor solutions across samples and adds to what is known about the psychometric properties of the RBANS.

Additionally, the present study has empirically supported the position that differences in RBANS factor solutions are primarily due to methodological decisions and are not solely related to unique sample characteristics. Simply put, the RBANS factor structure is relatively invariant across diverse samples. Previous RBANS factor-analytic studies (Duff et al., [19]; Garcia et al., [24]; King et al., [39]; Schmitt et al., [18]; Wilde, [67]) revealed subtle differences between solutions and suggested they were due to sample differences. However, these findings indicate that the underlying structure does not meaningfully differ across samples when empirically supported factor-retention guidelines are considered. Similar methods could be implemented to help clinicians and researchers understand the psychometric properties of other neuropsychological measures when the literature includes divergent factor-analytic results.

This investigation of the RBANS provides important clinical insights. Consistent with previous findings (Carlozzi et al., [10]; Duff et al., [19]; Garcia et al., [24]; King et al., [39]; Schmitt et al., [18]; Wilde, [67]), the underlying structure of the RBANS is inconsistent with the theoretical index scores. The RBANS component structure suggests the memory and visuospatial constructs are most reliably assessed. Furthermore, the Picture Naming subtest demonstrated a ceiling effect in clinical and nonclinical samples, thus impacting overall clinical utility. Also noteworthy, the Digit Span subtest did not converge with other tasks within the RBANS. This information in combination with findings that select index scores have problematic reliabilities, suggests that clinicians should be cautious when interpreting those composite scores. It seems that the RBANS is an appropriate measure to use when the purpose of the evaluation is to efficiently and effectively evaluate memory and visuospatial abilities. If a more comprehensive or refined neuropsychological picture is sought, a clinician is encouraged to augment the RBANS by administering measures sensitive to language, attention, and other relevant neurocognitive constructs.

Acknowledgments

The authors wish to thank Kevin Duff, Ph.D., and Mark C. Wilde, Ph.D., for providing correlational matrixes for reanalysis.

References

- Aupperle, R. L., Beatty, W. W., Shelton, F. N., & Gontkovsky, S. T. (2002). Three screening batteries to detect cognitive impairment in multiple sclerosis. *Multiple Sclerosis, 8*, 382–389. doi:10.1191/1352458502ms832oa
- Barrett, P. T. (1986). Factor comparison: An examination of three methods. *Personality and Individual Differences, 7*, 327–340. doi:10.1016/0191-8869(86)90008-5
- Barrett, P. T. (2005). Orthosim 2.1 (Version 2.1): Target comparison matrix fitting [computer software and manual]. Retrieved from <http://www.pbarrett.net/orthosim/orthosim.html>
- Barrett, P. T., Petrides, K. V., Eysenck, S. B. G., & Eysenck, H. J. (1998). The Eysenck Personality Questionnaire: An examination of the factorial similarity of P, E, N, and L across 34 countries. *Personality and Individual Differences, 25*, 805–819. doi:10.1016/s0191-8869(98)00026-9
- Beatty, W. W. (2004). RBANS analysis of verbal memory in multiple sclerosis. *Archives of Clinical Neuropsychology, 19*, 825–834. doi:10.1016/j.acn.2003.12.001
- Beatty, W. W., Ryder, K. A., Gontkovsky, S. T., Scott, J. G., McSwan, K. L., & Bharucha, K. J. (2003). Analyzing the subcortical dementia syndrome of Parkinson's disease using the RBANS. *Archives of Clinical Neuropsychology, 18*, 509–520. doi:10.1016/s0887-6177(02)00148-8
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588–606. doi:10.1037/0033-2909.88.3.588
- Calamia, M., Markon, K., Denburg, N. L., & Tranel, D. (2011). Developing a short form of Benton's Judgment of Line Orientation Test: An item response theory approach. *The Clinical Neuropsychologist, 25*, 670–684. doi:10.1080/13854046.2011.564209
- Camara, W. J., Nathan, J. S., & Puente, A. E. (2000). Psychological test usage: Implications in professional psychology. *Professional Psychology: Research and Practice, 31*, 141–154. doi:10.1037//0735-7028.31.2.141
- Carlozzi, N. E., Horner, M. D., Yang, C., & Tilley, B. C. (2008). Factor analysis of the Repeatable Battery for the Assessment of Neuropsychological Status. *Applied Neuropsychology, 15*, 274–279. doi:10.1080/09084280802325124
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York, NY: Cambridge University Press.
- Cattell, R. B. (1966). The screen test for the number of factors. *Multivariate Behavioral Research, 1*, 245–276. doi:10.1207/s15327906mbr0102_10
- Church, A. T., & Burke, P. J. (1994). Exploratory and confirmatory tests of the Big Five and Tellegen's three- and four-dimensional models. *Journal of Personality and Social Psychology, 66*, 93–114. doi:10.1037/0022-3514.66.1.93
- Crawford, A. V., Green, S. B., Lo, W. J., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Measurement, 70*, 885–901. doi:10.1177/0013164410379332
- Deary, I. J. (2000). *Looking down on human intelligence: From psychometrics to the brain*. Oxford, UK: Oxford University Press.
- Delis, D. C., Jacobson, M., Bondi, M. W., Hamilton, J. M., & Salmon, D. P. (2003). The myth of testing construct validity using factor analysis or correlations with normal or mixed clinical populations: Lessons from memory assessment. *Journal of the International Neuropsychological Society, 9*, 936–946. doi:10.1017/s1355617703960139
- Dowling, N. M., Hermann, B., La Rue, A., & Sager, M. A. (2010). Latent structure and factorial invariance of a neuropsychological test battery for the study of preclinical Alzheimer's disease. *Neuropsychology, 24*, 742–756. doi:10.1037/a0020176

- Duff, K. D., Hobson, V. L., Beglinger, L. J., & O'Bryant, S. E. (2010). Diagnostic accuracy of the RBANS in mild cognitive impairment: Limitations on assessing milder impairments. *Archives of Clinical Neuropsychology*, 25, 429–441. doi:10.1093/arclin/acq045
- Duff, K. D., Langbehn, D. R., Schoenberg, M. R., Moser, D. J., Baade, L. E., Mold, J. W., ... Adams, R. L. (2006). Examining the Repeatable Battery for the Assessment of Neuropsychological Status: Factor analytic studies in an elderly sample. *American Journal of Geriatric Psychiatry*, 14, 976–979. doi:10.1097/01.jgp.0000229690.70011.cd
- Duff, K. D., Langbehn, D. R., Schoenberg, M. R., Moser, D. J., Baade, L. E., Mold, J. W., ... Adams, R. L. (2009). Normative data on and psychometric properties of verbal and visual indexes of the RBANS in older adults. *The Clinical Neuropsychologist*, 23, 39–50. doi:10.1080/13854040701861391
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272–299. doi:10.1037/1082-989X.4.3.272
- Fava, J. L., & Velicer, W. F. (1992a). The effects of overextraction on factor and component analysis. *Multivariate Behavioral Research*, 27, 387–415. doi:10.1207/s15327906mbr2703_5
- Fava, J. L., & Velicer, W. F. (1992b). An empirical comparison of factor, image, component, and scale scores. *Multivariate Behavioral Research*, 27, 301–322. doi:10.1207/s15327906mbr2703_1
- Garcia, C., Leahy, B., Corradi, K., & Forchetti, C. (2008). Component structure of the Repeatable Battery for the Assessment of Neuropsychological Status in dementia. *Archives of Clinical Neuropsychology*, 23, 63–72. doi:10.1016/j.acn.2007.08.008
- Gignac, G. E., Bates, T. C., & Jang, K. (2007). Implications relevant to CFA model misfit, reliability, and the five factor model as measured by the NEO-FFI. *Personality and Individual Differences*, 43, 1051–1062. doi:10.1016/j.paid.2007.02.024
- Goldberg, L. R., & Velicer, W. F. (2006). Principles of exploratory factor analysis. In S. Strack (Ed.), *Differentiating normal and abnormal personality* (2nd ed., pp. 209–237). New York, NY: Springer.
- Grice, J. W. (2001a). A comparison of factor scores under conditions of factor obliquity. *Psychological Methods*, 6, 67–83. doi:10.1037/1082-989X.6.1.67
- Grice, J. W. (2001b). Computing and evaluating factor scores. *Psychological Methods*, 6, 430–450. doi:10.1037//1082-989X.6.4.430
- Hobart, M. P., Goldberg, R., Bartko, J. J., & Gold, J. M. (1999). Repeatable Battery for the Assessment of Neuropsychological Status as a screening test in schizophrenia. II: Convergent/discriminant validity and diagnostic group comparisons. *American Journal of Psychiatry*, 156, 1951–1957.
- Hoelzle, J. B. (2008). *Neuropsychological assessment in the Cattell-Horn-Carroll (CHC) cognitive abilities model* (Unpublished doctoral dissertation). The University of Toledo, Toledo, OH.
- Hoelzle, J. B., & Meyer, G. J. (2009). The invariant component structure of the Personality Assessment Inventory (PAI) full scales. *Journal of Personality Assessment*, 91, 175–186. doi:10.1080/00223890802634316
- Hoelzle, J. B., & Meyer, G. J. (2013). Exploratory factor analysis: Basics and beyond. In I. B. Weiner, J. A. Schinka, & W. F. Velicer (Eds.), *Handbook of psychology: Research methods in psychology* (2nd ed., pp. 164–188). Hoboken, NJ: John Wiley & Sons.
- Hoelzle, J. B., Nelson, N. W., & Smith, C. A. (2011). Comparison of the Wechsler Memory Scale-Fourth Edition (WMS-IV) and Third Edition (WMS-III) dimensional structures: Improved ability to evaluate auditory and visual constructs. *Journal of Clinical and Experimental Neuropsychology*, 33, 283–291. doi:10.1080/13803395.2010.511603

- Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review*, 14, 332–346.
doi:10.1177/1088868310361240
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 19, 179–185. doi:10.1007/BF02289447
- Hubbard, R., & Allen, S. J. (1987). An empirical comparison of alternative methods for principal component extraction. *Journal of Business Research*, 15, 173–190. doi:10.1016/0148-2963(84)90047-X
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, 141–151. doi:10.1077/01316446002000116
- Kaiser, H. F. (1981). A revised measure of sampling adequacy for factor-analytic data matrices. *Educational and Psychological Measurement*, 41, 379–381. doi:10.1177/001316448104100216
- King, L. C., Bailie, J. M., Kinney, D. I., & Nitch, S. R. (2012). Is the Repeatable Battery for the Assessment of Neuropsychological Status factor structure appropriate for inpatient psychiatry? An exploratory and higher-order analysis. *Archives of Clinical Neuropsychology*, 27, 756–765.
doi:10.1093/arclin/acs062
- Kokmen, E., Naessens, J. M., & Offord, K. P. (1987). A short test of mental status: Description and preliminary results. *Mayo Clinic Proceedings*, 62, 281–288. doi:10.1016/S0025-6196(12)61905-3
- Korth, B., & Tucker, L. R. (1975). The distribution of chance congruence coefficients from simulated data. *Psychometrika*, 40, 361–372. doi:10.1007/BF02291763
- Larrabee, G. J. (2003). Lessons on measuring construct validity: A commentary on Delis, Jacobson, Bondi, Hamilton, and Salmon. *Journal of the International Neuropsychological Society*, 9, 947–953. doi:10.1017/S1355617703960140
- Larson, E. B., Kirschner, K., Bode, R. K., Heinemann, A., & Goodman, R. (2005). Construct and predictive validity of the Repeatable Battery for the Assessment of Neuropsychological Status in the evaluation of stroke patients. *Journal of Clinical and Experimental Neuropsychology*, 27, 16–32.
doi:10.1080/138033990513564
- Lee, K., & Ashton, M. C. (2007). Factor analysis in personality research. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 424–443). New York, NY: Guilford.
- Levine, M. S. (1977). *Canonical analysis and factor comparison*. Beverly Hills, CA: Sage.
- Linn, R. L. (1968). A Monte Carlo approach to the number of factors problem. *Psychometrika*, 33, 37–71. doi:10.1007/bf02289675
- Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 2, 57–64. doi:10.1027/1614-2241.2.2.57
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4, 84–99. doi:10.1037/1082-989X.4.1.84
- McCrae, R. R., Zonderman, A. B., Costa, P. T., Jr., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the revised NEO Personality Inventory: Confirmatory factor analysis versus Procrustes rotation. *Journal of Personality and Social Psychology*, 70, 552–566.
doi:10.1037/0022-3514.70.3.552
- McKay, C., Casey, J. E., Wertheimer, J., & Fichtenberg, N. L. (2007). Reliability and validity of the RBANS in a traumatic brain injured sample. *Archives of Clinical Neuropsychology*, 22, 228–241.
doi:10.1016/j.bbr.2011.03.031

- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performance as scientific inquiry into scoring meaning. *American Psychologist*, 9, 741–749. doi:10.1037/0003-066x.50.9.741
- Morey, L. C. (1991). *Personality Assessment Inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Nelson, N. W., Sweet, J. J., Berry, D. T. R., Bryant, F. B., & Grancher, R. P. (2007). Response validity in forensic neuropsychology: Exploratory factor analytic evidence of distinct cognitive and psychological constructs. *Journal of the International Neuropsychological Society*, 13, 440–449. doi:10.1017/S1355617707070373
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavioral Research Methods, Instruments & Computers*, 32, 396–402. doi:10.3758/bf03200807
- O'Connor, B. P. (2002). The search for dimensional structure differences between normality and abnormality: A statistical review of published data on personality and psychopathology. *Journal of Personality and Social Psychology*, 83, 962–982. doi:10.1037/0022-3514.83.4.962
- Preacher, K. J., & MacCallum, R. C. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics*, 2, 13–43. doi:10.1207/S15328031US0201_02
- Psychological Corporation. (2001). *Manual for the Wechsler Test of Adult Reading (WTAR)*. San Antonio, TX: Psychological Corporation.
- Randolph, C. (1998). *Repeatable Battery for the Assessment of Neuropsychological Status*. Bloomington, MN: NSC Pearson.
- Randolph, C. (2012). *Repeatable Battery for the Assessment of Neuropsychological Status Update (RBANS Update)*. Bloomington, MN: NSC Pearson.
- Schmitt, A. L., Livingston, R. B., Smernoff, E. N., Reese, E. M., Hafer, D. G., & Harris, J. B. (2010). Factor analysis of the Repeatable Battery for the Assessment of Neuropsychology Status (RBANS) in a large sample of patients suspected of dementia. *Applied Neuropsychology*, 17, 8–17. doi:10.1080/09084280903297719
- Schoenberg, M. R., & Scott, J. G. (2011). The neuropsychology referral and answering the referral question. *The little black book of neuropsychology: A syndrome-based approach*. New York, NY: Springer.
- Tabachnick, B. G., & Fidell, L. S. (2013). Principal components and factor analysis. In B. G. Tabachnick & L. S. Fidell (Eds.), *Using multivariate statistics* (6th ed., pp. 612–680). Boston, MA: Pearson Education.
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41, 321–327. doi:10.1007/BF02293557
- Velicer, W. F., & Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, 3, 231–251. doi:10.1037/1082-989x.3.2.231
- Wechsler, D. (2009a). *Advanced Clinical Solutions (ACS) for the Wechsler Adult Intelligence Scale-Fourth Edition and Wechsler Memory Scale-Fourth Edition*. San Antonio, TX: Pearson Assessments.
- Wechsler, D. (2009b). *Wechsler Adult Intelligence Scale-Fourth Edition*. San Antonio, TX: Pearson Assessments.
- Wilde, M. C. (2006). The validity of the Repeatable Battery of Neuropsychological Status in acute stroke. *The Clinical Neuropsychologist*, 20, 702–715. doi:10.1080/13854040500246901
- Wilk, C. M., Gold, J. M., Humber, K., Dickerson, F., Fenton, W. S., & Buchanan, R. W. (2004). Brief cognitive assessment in schizophrenia: Normative data for the Repeatable Battery for the

Assessment of Neuropsychological Status. *Schizophrenia Research*, 70, 175–186.
doi:10.1016/j.schres.2003.10.009

Woodhouse, J., Heyanka, D. J., Scott, J., Vincent, A., Roebuck-Spencer, T., Domboski-Davidson, K., ... Adams, R. (2013). Efficacy of the ANAM General Neuropsychological Screening Battery (ANAM GNS) for detecting neurocognitive impairment in a mixed clinical sample. *The Clinical Neuropsychologist*, 27, 376–385. doi:10.1080/13854046.2012.762427

Zwick, W. R., & Velicer, W. F. (1982). Factors influencing four rules for determining the number of components to retain. *Multivariate Behavioral Research*, 17, 253–269.
doi:10.1207/s15327906mbr1702_5

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules to determine the number of components to retain. *Psychological Bulletin*, 99, 432–442. doi:10.1037/0033-2909.99.3.432