

## Factor structure of the CERAD neuropsychological battery

MILTON E. STRAUSS<sup>1,2</sup> AND THOMAS FRITSCH<sup>2,3</sup>

<sup>1</sup>Department of Psychology, Case Western Reserve University, Cleveland, Ohio

<sup>2</sup>University Memory and Aging Center of University Hospitals of Cleveland and Case Western Reserve University, Cleveland, Ohio

<sup>3</sup>Department of Neurology, Case Western Reserve University School of Medicine, Cleveland, Ohio

(RECEIVED August 25, 2003; REVISED December 9, 2003; ACCEPTED December 22, 2003)

### Abstract

The Consortium to Establish a Registry for Alzheimer's Disease (CERAD) neuropsychological battery was developed to evaluate cognitive impairments associated with Alzheimer's disease (AD). Previous studies have suggested that the battery is multi-dimensional, represented by either 3 or 5 dimensions. In this study a principal factor analysis was conducted using contemporary quantitative methods for determining the number of factors. Exploratory factor analysis of the CERAD battery and MMSE was conducted using one-half of the CERAD database (total  $N = 969$ ). Glorfeld's modification of Horn's parallel analysis method suggested that there was 1 common factor in the variable matrix. Characterization of patterns of deficits in AD requires supplementation of measures derived from the CERAD and MMSE with other tests. (*JINS*, 2004, *10*, 559–565.)

**Keywords:** Alzheimer's disease, CERAD, Factor analysis, Neuropsychological assessment

### INTRODUCTION

An important contribution of the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) was the standardization of a brief neuropsychological battery (CERAD NP) for use in the diagnosis of Alzheimer's disease (AD) and the characterization of severity of cognitive impairment (Morris et al., 1993; Welsh-Bohmer & Mohs, 1997). The basic battery consists of six measures in addition to the Mini-Mental State Examination (MMSE, Folstein et al., 1975): (1) total score on a modified Boston Naming Test; (2) number of unique responses on a verbal fluency test; (3) accuracy of design copy (constructional praxis); and three measures derived from a word-list learning test; (4) immediate recall; (5) delayed recall; and (6) recognition memory.<sup>1</sup> The battery is sensitive to early-stage dementia, reliable over a 1-month period, and sensitive to change over longer time spans (Fillenbaum et al., 2002; Morris et al., 1989;

1993; Welsh et al., 1991; 1992). There are now more than 100 published studies of clinically normal individuals or persons with dementia in which the CERAD NP is used to characterize neuropsychological function, as well as studies of other patient groups, such as schizophrenia (McGurk et al., 2000), depression (Paradiso et al., 1997), and anxiety (Xavier et al., 2001).

The CERAD NP battery was proposed as assessing three aspects of cognitive performance considered to be relatively independent by the developers: language, memory, and praxis (Morris et al., 1989). In support of this concept, Morris et al. (1989) reported an analysis of the covariation among the measures generally approximating these domains. However, that analysis is difficult to interpret because the method of factor extraction, criteria for determining the number of factors, and rotation were not specified. There appears to be only one other evaluation of the constructs represented in the CERAD NP. Collie et al. (1999) studied healthy older adults and analyzed subsets of items. They reported that five principal components with Eigenvalues greater than 1.0, accounting for approximately 63% of the total variance, appeared to best represent the CERAD battery.

The differences between the findings of these two studies may reflect sample differences, the measures included

Reprint requests to: Milton E. Strauss, Ph.D., University Memory and Aging Center, 12200 Fairhill Road, Cleveland, OH 44120. E-mail: milton.strauss@cwru.edu

<sup>1</sup>The Trail Making Test was added later in the course of the CERAD project but was not available for most subjects and so was not included in this study.

in the analysis, and, perhaps, factor extraction methods. Morris et al. (1989) treated the MMSE as a single measure, while Collie et al. (1999) represented the MMSE in five constituent components (attention, language, orientation, recall, and registration). Morris et al. (1989) included three memory measures from the list-learning task (memory over three trials, delayed recall, and an adjusted correct recognition score), while Collie et al. (1999) included six; three immediate memory trials, recall, recognition hits, and recognition correct rejections.

Specific differences aside, both studies appear to suggest that the CERAD NP is a multidimensional test battery. However, such a conclusion may be premature because of at least three aspects of the methods and procedures in these two studies.

A purpose of both of these studies was the identification of the psychological constructs represented in the CERAD NP. Collie et al. (1999) used principal components analysis (PCA) to achieve this and we assume this for Morris et al. (1989) too, because of the widespread popularity of PCA (Floyd & Widaman, 1995). However, principal components do not represent the latent variables (constructs) underlying relations among a set of measures in a sample. Principal components analysis is a data reduction method, a mathematical simplification of a correlation/covariance matrix. It does not differentiate between reliable variance and error variance, or between variance shared by two or more tests (common variance) and reliable variance unique to each measure (specific variance, see Rummel, 1970).

Contemporary methodologists recommend the use of exploratory factor analysis (EFA) to identify the dimensions underlying the covariation in a set of measures (Floyd & Widaman, 1995; Gorsuch, 1983; Tabachnik & Fidell, 1996), especially when the number of variables is small and communalities are not uniformly high (Gorsuch, 1983). Exploratory factor analysis considers only reliable variance, and factor loadings are estimated from only the variance shared by at least two variables. Exploratory factor analysis is a latent trait analysis, and so EFA loadings may be thought of as regression weights that predict observed scores from the estimated latent constructs (Floyd & Widaman, 1995).

A second matter in the analysis of correlation or covariance structures is the number of factors to be retained for rotation. Two widely used criteria are the number of Eigenvalues greater than 1 and the scree test, which retains the number of factors prior to a "bend" in the plot of Eigenvalues against factor number. Recent studies using Monte Carlo simulations and more quantitative methods of determining the number of factors in a matrix indicate that the Eigenvalue criterion is much too liberal, leading to overestimation of the number of valid factors. Scree is fairly accurate, though less so than more quantitative methods, perhaps because of the inherent subjectivity of visual inspection judgments (Zwick & Velicer, 1986). The factor extraction method and factor retention criteria is likely to have led to an overestimation of the number of dimensions described for the CERAD battery by Morris et al. (1989) and Collie et al. (1999).

Multiple measures were derived from single tests in both studies, which confounds method and trait (construct) variance (Campbell & Fiske, 1959). Thus, a large verbal memory factor would be expected when several (3 for Morris et al., 1989, and 6 for Collie et al., 1999) indices from one test are included in a single analysis. The separation of method from construct variance requires that there be multiple indicators from independent measurements as converging operations (Garner et al., 1956); without these, method factors are highly likely (Strauss et al., 2000).

The purpose of the present study was to identify the latent trait structure of the CERAD NP battery in the consortium's database of over 900 patients. The overarching hypothesis was that shared variance in the battery would be accounted for by a single factor. As a brief battery, the CERAD does not contain multiple indicators of constructs, and so constructs are likely to be under-specified (Bollen, 1989). Additionally, multiple factors are not invariable in groups with generalized impairment, such as in dementia, even when several indicators of presumably distinct constructs (e.g., memory, visual-spatial ability, executive functions) are included in the assessment (Strauss & Summerfelt, 2002).

## METHODS

### Research Participants

The data in this study were derived from the CD of the database developed by the Consortium to Establish a Registry for Alzheimer's Disease (Morris et al., 1989). The database includes standardized neuropsychological test results for participants with AD and normal controls at 20 different sites across the United States. Data include entry visits and such annual follow-up visits as were completed. For this study, we drew all participants who met NINCDS-ADRDA criteria (McKhann et al., 1984) for probable or possible AD at entry into the registry, had Clinical Dementia Rating (CDR; Hughes et al., 1982) scores of 1 (*mild*) or 2 (*moderate*), and neuropsychological data at the initial visit. This resulted in a sample size of 969 persons.

### Measures

The CERAD NP (Morris et al., 1993) battery consists of the following tests: (1) Modified Boston Naming Test, which measures Confrontation Naming (range of scores: 0–15); (2) Verbal Fluency, which measures verbal production ability, semantic memory, and language (range of scores: zero–high); (3) Word List Memory, a test of verbal memory (range of scores 0–30); (4) Constructional Praxis, which measures visuospatial ability (range of scores 0–11); (5) Word List Recall, which measures delayed verbal memory (range of scores: 0–10); and (6) Word List Recognition, which also measures delayed verbal memory (range of scores: 0–20).

The Mini-Mental State Exam (MMSE, Folstein et al., 1975) was also administered. Following the approach of Collie and associates (1999) the MMSE was scored as five subscales: (1) *Orientation*, consisting of questions about the year, season, date, day, month, state, county, town, hospital, and floor; (2) *Registration*, requiring the subject to repeat the names of three objects; (3) *Attention*, spelling “world” backwards; (4) *Recall*, consisting of memory for three objects presented earlier in the test; and (5) *Language/Praxis*, consisting of naming two objects, repeating a phrase spoken by the examiner, following a three-stage command, reading and obeying a command, writing a sentence, and copying a simple geometric design. Jones and Gallo’s (2000) factor analysis of the MMSE in a large community sample supports this segregation of items, first suggested by Folstein et al. (1975).

## Methods of Analysis

The CERAD sample was randomly divided into a derivation sample ( $N = 500$ ) and a validation sample ( $N = 469$ ) of cases, using only subjects for whom there were complete demographic data on the first visit. Since correlations are a function of score distributions, the frequency distributions of all measures were first examined for skewness.

We determined the number of factors to retain in the factor analysis using Glorfeld’s (1995) modification of the Horn (1965) parallel analysis model. One-hundred samples of random number matrices were generated, each of which consisted of  $N$  cases and  $k$  variables, where  $N$  and  $k$  are the number of subjects in a sample and the number of tests included in the actual data set, respectively. A principal components analysis was computed for each random number correlation matrix. Since the simulated data are random, the Eigenvalues reflect chance associations. The number of factors retained in the factor analysis was the number for which the Eigenvalues in the CERAD NP analysis exceeded the 95th percentile of the random Eigenvalue distribution. All analyses were conducted using SPSS 10.1 using syntax by O’Connor (2000).

## RESULTS

### Characteristics of Participants

All subjects met criteria established by the National Institute on Neurological and Communicative Disorders and the Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRDA; McKhann et al., 1984) for the diagnosis of probable (90.5%) or possible AD (9.5%). As a part of their comprehensive CERAD evaluation, they received a medical examination, neuropsychological testing, functional screening (Blessed Dementia Rating Scale), and laboratory and neurological examinations. Many subjects had CT or MRI scans of the brain. To be eligible for CERAD, subjects had to be free of major co-morbid medical problems, such as cancer, cardiac or respiratory disease, hyper-

tension, or major depression or psychiatric illness. The demographic and clinical characteristics of the participants in the derivation and validation samples are summarized in Table 1. The mean age of subjects at entry was 72.8 years ( $SD = 8.0$ ); mean education was 12.5 years ( $SD = 1.6$ ); 58.9% were female; 81.5% were White and 18.5% were minorities (mostly African American). According to Clinical Dementia Rating (CDR) Scale staging (Hughes et al., 1982), 57.6% had mild dementia ( $CDR = 1$ ) and 42.4% had moderate dementia ( $CDR = 2$ ). The participants in the two samples did not differ for the demographic and clinical variables summarized in Table 1 ( $ps > .20$ ).

### Neuropsychological Test Performance

The performances of each group on each of the CERAD NP variables considered for analysis are summarized in Table 2. CERAD NP batteries were unavailable for 2 subjects and between 2 and 10 other subjects were missing at least one test score. Consequently, the factor analyses were based on the 475 of 500 subjects in the derivation sample with complete NP data. The cross-validation was conducted using the 438 remaining complete records.

A number of the variables had skewed distributions, as suggested by scores  $\pm 2$  standard deviations from the mean being outside of the permissible range. Three measures had extremely skewed distributions. Eighty percent of the sample had perfect scores on MMSE Registration; 71% had scores of zero for MMSE Recall; and 61.6% had scores of zero on delayed recall for the list learning task. Because of these extreme distributions, these measures were not included in subsequent analyses. As shown in Table 2, the scores of the validation sample did not differ from those in the derivation sample ( $ps > .14$ ).

**Table 1.** Demographic and clinical characteristics of participants in the derivation and validation samples

Variable	Derivation sample ( $N = 500$ )	Validation sample ( $N = 469$ )
Mean age ( $SD$ )	72.87 (8.20)	72.63 (7.85)
Mean education ( $SD$ )	12.69 (6.71)	12.29 (5.44)
Percent female ( $N$ )	57.8 (289)	60.1 (282)
Ethnicity		
Percent White ( $N$ )	80.0 (400)	83.2 (390)
Percent minority ( $N$ )	20.0 (100)	16.8 (79)
Diagnosis		
Percent with “probable” AD ( $N$ )	90.2 (451)	90.8 (426)
Percent with “possible” AD ( $N$ )	9.8 (49)	9.2 (43)
CDR		
Percent with “mild” AD ( $N$ )	58.2 (291)	56.9 (267)
Percent with “moderate” AD ( $N$ )	41.8 (209)	43.1 (202)
Mean Blessed ADL ( $SD$ )	4.53 (2.42)	4.50 (2.31)

*Note.*  $p$  values are derived from  $t$  tests (homogenous variances) and chi-square tests, as appropriate.

**Table 2.** Neuropsychological test performance of participants in the derivation and validation samples

Variable	Derivation sample			Validation sample			<i>p</i> value
	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	
MMSE							
Orientation	5.10	2.50	493	4.93	2.57	461	.29
Attention	3.01	1.87	496	2.89	1.90	465	.31
Registration	2.70	0.66	496	2.65	0.72	465	.28
Recall	0.40	0.70	496	0.41	0.73	465	.87
Language	6.08	1.71	490	5.91	1.76	464	.14
Praxis	0.38	0.49	494	0.40	0.49	466	.43
Naming	10.25	3.47	498	10.57	3.54	469	.16
Fluency	7.52	3.99	498	7.46	4.26	468	.84
Praxis	6.92	2.79	495	6.84	2.84	462	.69
List learning							
Immediate recall	7.54	4.41	494	7.29	4.58	465	.39
Delayed recall	0.73	1.26	490	0.68	1.20	463	.46
Recognition	13.90	3.60	488	13.74	3.68	456	.51

Note. *p* values are derived from *t* tests.

## Factor Analysis

### Derivation sample

Based on the evaluation of the score distributions the following variables were included in the factor analysis: Orientation, Attention, and Language/Praxis from the MMSE, and Total Naming, Verbal Fluency, Constructional Praxis, and a composite Verbal Memory measure, from the CERAD. The composite consisted of the average standardized immediate memory and standardized recognition score on the CERAD list-learning task. This composite was used instead of each component in order to avoid a test-specific factor. The two measures correlated with an  $r = .41$ ,  $N = 475$ ,  $p < .001$ .

The seven retained CERAD NP scores were subjected to principal factor analysis of the correlation matrix, which standardized all measures to a common metric. Communalities were estimated by squared multiple correlations. The first factor had an Eigenvalue of 3.471 and the value of the second was .995. The 95th percentile of the distribution of random Eigenvalues were 1.229 and 1.146, for the first and second factors respectively. By both the parallel analysis criterion and the more liberal Eigenvalue greater than 1.0 criterion, only one factor was necessary to represent the constructs underlying the shared variance among these measures.

The factor loadings and communalities in the derivation sample are presented in the first two columns of Table 3. The communality is the proportion of variance shared by an observed score and an underlying construct, which in the case of a one-factor solution is the squared factor loading. As may be seen in the table, each of the NP measures, with the exception of constructional praxis, has good loadings on this factor. As the communality estimates indicate, the

factor accounts for from one-third to over one-half of the reliable variance in each of the other variables.

Principal components analyses of 100 random data sets were computed for the validation sample, as well. Eigenvalues at the 95th percentile were 1.253 and 1.142 for the first two components. As was found in the derivation sample, only one factor had an Eigenvalue greater than 1.0. This Eigenvalue was 3.689; that for the second factor was .871. These values are similar to those in the derivation sample. The factor loadings and communalities of the seven measures are presented in the last two columns of Table 3. Inspection suggests substantial similarity in the loading patterns. Excellent replication of the solution is suggested by the coefficient of congruence between these two factors, which is .998. The coefficient of congruence is an estimate of the correlation between the factor scores (Gorsuch, 1983).

The common factor accounts for 41–45% of the reliable variance in the two samples. As Kaufman (1990) describes,

**Table 3.** Factor structure of CERAD NP measures in derivation and validation samples

Test	Derivation sample		Validation sample	
	Factor loading	Communality	Factor loading	Communality
Orientation	.60	.36	.60	.36
Attention	.71	.51	.66	.44
Language/praxis	.75	.56	.74	.55
Naming	.61	.37	.69	.48
Fluency	.68	.46	.73	.53
Praxis	.48	.24	.56	.31
Memory	.65	.42	.68	.46
% Variance	41.53		45.02	



**Table 4.** Estimates of specific reliable variance in CERAD NP measures

Variance	Measure					
	Orientation	Language	Naming	Fluency	Praxis	Memory
Reliable variance	.72	.64	.77	.60	.64	.58
Common variance	.36	.56	.37	.46	.24	.42
Specific variance	.36	.08	.40	.14	.40	.16
Error variance	.28	.36	.23	.40	.36	.42

*Note.* Reliability coefficients (coefficient alpha) were based on CERAD data with the exception of Naming, which was computed from 250 cases in the University Memory and Aging Center. Attention is omitted since reliability could not be estimated. See text for further explanation.

reliable variance consists of common variance, which is indicated by squared factor loading of a measure, and specific variance. The latter reflects constructs or traits that are not shared with other measures in a test battery. Specific variance may be estimated by the difference between the communality and the reliability of the test. We estimated reliability using Chronbach’s alpha for each variable in the analysis except MMSE attention, since it is a single item measure. Reliability was estimated in somewhat different ways for these measures. Coefficient alpha was estimated across items for Orientation and Language. The total score for verbal fluency is the sum of the number of responses in three successive 15-s intervals. We estimated coefficient alpha based on these three intervals. The reliability of Memory was estimated from the two components of this score. The CERAD database provided only the total score for Naming and so reliability was estimated in a sample of 250 probable and possible AD cases in our local database. In our database we record the number of correct responses for high-frequency, medium-frequency, and low-frequency words separately. We estimated coefficient alpha on the basis of these three sub-scores. The reliability coefficients (Table 4) are moderate (.58) to good (.72), particularly considering the brevity of each measure.

**Table 5.** Principal factor analysis loadings with two praxis tests

Test	Derivation sample		Validation sample	
	Factor		Factor	
	I	II	I	II
Orientation	.63	.12	.62	.13
Attention	.52	.51	.51	.43
Language	.61	.32	.63	.26
Naming	.59	.19	.63	.27
Fluency	.65	.21	.71	.23
Praxis	.16	.82	.26	.82
Memory	.71	.11	.71	.16
MMSE Praxis	.16	.61	.16	.63
% variance	29.6	19.0	31.9	18.6

*Note.* Italicized indicates most salient loadings on each factor.

Table 4 displays the estimated variance components of the CERAD measures other than MMSE Attention, for which reliability could not be estimated. Comparing the specific variance of a variable to its error variance can serve to evaluate the extent to which other constructs might be tapped by the test (Kaufman, 1990). The orientation measure from the MMSE, naming, and praxis each have substantial specific variance, larger in each case than the error variance for the measure, suggesting that each assesses a construct in addition to general dementia severity. Indeed, the specific variance of the praxis score is greater than the variance associated with the general severity factor. This test had the lowest factor loading of the six scores. When the praxis item from the MMSE language/praxis score was included as a separate variable, a second factor did emerge in the analysis (Table 5). The first factor accounts for 30–32% of the reliable variance, while the second factor, defined mainly by the two praxis measures, accounts for an additional 19% of reliable variance in the system.

To evaluate the interpretation of the factor as a severity index, we computed a factor score for each subject. The variables were first standardized as *z* scores and then summed, using both factor scores and unit weighting. As would be expected for a severity index, there was a substantial difference between the standardized factor score for CDR mild cases ( $N = 536, M \pm SD = .36 \pm .74$  and  $.31 \pm .55$ , for factor and for unit weighted scores, respectively) and the CDR moderate cases ( $N = 377, M \pm SD = -.52 \pm .90$  and  $-.36 \pm .67$ , for factor and for unit weighted scores, respectively;  $t_s > 15, p_s < .001$ ). Since the distinction between probable and possible AD is based on an etiological hypothesis rather than severity of impairment, a difference between groups for either method of computing a difference score was not expected ( $M_{\text{difference}} < .08, t_s < 1$ ).<sup>2</sup>

## DISCUSSION

The replicated results of principal factor analysis of the CERAD measures that were suitable for factoring suggests

<sup>2</sup>A factor analysis computed using only probable AD cases produced comparable results to the factor analysis reported here.

that one latent variable of dementia severity accounts for the shared variance among them and that the components of the CERAD battery, in conjunction with the MMSE, should not themselves be used to evaluate potential patterns of deficits within AD. The specific reliable variance of three measures, the language subscale of the MMSE, the CERAD fluency and CERAD memory composite scores, is substantially smaller than the error variance of each, which suggests that these measures assess only the overall severity of dementia. Orientation, naming and praxis, on the other hand, have sufficient specific variance in relationship to error variance to suggest that these may be measuring other constructs. However, the specification of such additional constructs requires additional measures (Tabachnick & Fidell, 1996), as the analysis including an additional praxis measure demonstrated. Lowenstein et al.'s (2001) analysis of the more extensive NINCDS-ADRDA neuropsychological battery also found a multifactorial structure when multiple indicators of constructs were evaluated. This suggests that the CERAD battery be supplemented by additional measures of naming and praxis when this battery is used to describe qualitative features of cognition in AD.

Fisher et al. (1999) reported that there were three clusters or subtypes of AD patient as assessed by CERAD tests. Interestingly, memory scores did not differentiate among these types. The effective discriminators were naming, praxis, and verbal fluency. The first two of these were also found likely to have reasonable specific-construct variance in this factor analysis. This was not the case for verbal fluency, but this test was the weakest discriminator among patient subgroups in Fisher et al.'s (1999) report. One implication of these analyses, taken together is that supplementary measures of confrontation naming and praxis used in conjunction with the CERAD tests might be an efficient clinical research approach to making clearer discriminations among subgroups.

The results of the present analyses were influenced by a number of decisions about data reduction. We elected to use MMSE subscales to avoid the loading of a MMSE total score on multiple factors, should such have emerged. Fisher et al. (1999) did not include the MMSE in their cluster analysis. Several CERAD and MMSE variables were excluded because of highly skewed distributions, which violate assumptions of factor analysis (Tabachnick & Fidell, 1996). The results might have differed some had other criteria for retaining variables been used. The approach in this study minimizes the influence of method variance in the factor analysis on the NP battery and the results are strongly consistent with the hypothesis that a single construct is measured by the set of instruments. More differentiated assessment of cognitive functions in Alzheimer's disease would seem to require supplementary testing.

## ACKNOWLEDGMENTS

This research was supported in part by NIA Grant P50 AG08012 to Karl Herrup, Ph.D. We thank Jason Brandt, Marian Patterson, and Eric Youngstrom for helpful comments on this paper.

## REFERENCES

- Bollen, K.A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105.
- Collie, A., Shafiq-Antonacci, R., Maruff, P., Tyler, P., & Currie, J. (1999). Norms and the effects of demographic variables on a neuropsychological battery for use in healthy ageing Australian populations. *Australian and New Zealand Journal of Psychiatry*, *33*, 568–575.
- Fillenbaum, G.G., Unverzagt, F.W., Ganguli, M., Welsh-Bohmer, K.A., & Heyman, A. (2002). The CERAD Neuropsychological Battery: Performance of representative community and tertiary care samples of African-American and European-American elderly. In F.R. Ferraro (Ed.), *Minority and cross-cultural aspects of neuropsychological assessment* (pp. 45–77). Lisse, NL: Swets & Zeitlinger.
- Fisher, N.J., Rourke, B.P., & Bieliauskas, L.A. (1999). Neuropsychological subgroups of patients with Alzheimer's disease: An examination of the first 10 years of CERAD data. *Journal of Clinical and Experimental Neuropsychology*, *21*, 488–518.
- Floyd, F. & Widaman, K. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7*, 286–299.
- Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). 'Mini-Mental State': A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.
- Garner, W.R., Hake, H.W., & Eriksen, C.W. (1956). Operationism and the concept of perception. *Psychological Review*, *63*, 149–159.
- Glorfeld, L.W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement*, *55*, 377–393.
- Gorsuch, R.L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Horn, J.L. (1965). A rationale and test for the number of factors in a factor analysis. *Psychometrika*, *30*, 179–185.
- Hughes, C.P., Berg, L., Danziger, W., Coben, L.A., & Martin, R.L. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry*, *140*, 566–572.
- Jones, R.N. & Gallo, J.J. (2000). Dimensions of the Mini-Mental State Examination among community dwelling older adults. *Psychological Medicine*, *30*, 605–618.
- Kaufman, A. (1990). *Assessing adolescent and adult intelligence*. Boston: Allyn & Bacon.
- Lowenstein, D.A., Ownby, R., Schram, L., Acevedo, A., Rubert, M., & Argiuelles, T. (2001). An evaluation of the NINCDS-ADRDA neuropsychological criteria for the assessment of Alzheimer's disease: A confirmatory factor analysis of single versus multi-factor models. *Journal of Clinical and Experimental Neuropsychology*, *23*, 274–284.
- McGurk, S.R., Moriarity, P.J., Harvey, P.D., Parrella, M., White, L., & Davis, K.L. (2000). The longitudinal relationship of clinical symptoms, cognitive functioning, and adaptive life in geriatric schizophrenia. *Schizophrenia Research*, *42*, 47–55.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E.M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under

- the auspices of the Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology*, *34*, 939–944.
- Morris, J.C., Edland, S., Clark, C., Galasko, D., Koss, E., Mohs, R., van Belle, G., Fillenbaum, G., & Heyman, A. (1993). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part IV. Ratings of cognitive change in the longitudinal assessment of probable Alzheimer's disease. *Neurology*, *43*, 2457–2465.
- Morris, J.C., Heyman, A., Mohs, R.C., Hughes, J.P., van Belle, G., Fillenbaum, G., Mellits, E.D., & Clark, C. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, *39*, 1159–1165.
- O'Connor, B.P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments & Computers*, *32*, 396–402.
- Paradiso, S., Lambert, G.J., Garvey, M.J., & Robinson, R.G. (1997). Cognitive impairment in the euthymic phase of chronic unipolar depression. *Journal of Nervous and Mental Disease*, *185*, 748–754.
- Rummel, R.J. (1970). *Applied factor analysis*. Evanston, IL: Northwestern University Press.
- Strauss, M.E. & Summerfelt, A. (2002). The neuropsychological study of schizophrenia: A methodological perspective. In M.F. Lenzenweger & J.M. Hooley (Eds.), *Principles of experimental psychopathology: Essays in honor of Brendan A. Maher* (pp. 119–134). Washington, DC: American Psychological Association.
- Strauss, M.E., Thompson, P.T., Adams, N.L., Redline, S., & Burandt, C. (2000). Evaluation of a model of attention with confirmatory factor analysis. *Neuropsychology*, *14*, 201–208.
- Tabachnick, B.G. & Fidell, L.S. (1996). *Using multivariate statistics* (3rd ed.). Boston: Allyn & Bacon.
- Welsh, K.A., Butters, N., Hughes, J., Mohs, R., & Heyman, A. (1991). Detection of abnormal memory in mild cases of Alzheimer's disease using CERAD neuropsychological measures. *Archives of Neurology*, *48*, 278–281.
- Welsh, K.A., Butters, N., Hughes, J., Mohs, R., & Heyman, A. (1992). Detection and staging of dementia in Alzheimer's disease: Use of the neuropsychological measures developed for the Consortium to Establish a Registry for Alzheimer's Disease (CERAD). *Archives of Neurology*, *49*, 448–452.
- Welsh-Bohmer, K. & Mohs, R.C. (1997). Neuropsychological assessment of Alzheimer's disease. *Neurology*, *49*(Suppl. 3), S11–S13.
- Xavier, F.M., Ferraz, M.P., Trenti, C.M., Argimon, I., Bertolucci, P.H., Poyares, D., & Moriguchi, E.H. (2001). Generalized anxiety disorder in a population aged 80 years and older. *Revista de Saude Publica*, *35*, 294–302.
- Zwick, W.R. & Velicer, W.F. (1986). Comparisons of five rules for determining the number of components to retain. *Psychological Bulletin*, *99*, 432–442.