



Neuropsychological Assessment Battery™

Psychometric and Technical Manual

Travis White, PhD

Robert A. Stern, PhD

PAR Psychological Assessment Resources, Inc. • 16204 N. Florida Avenue • Lutz, FL 33549 • 1.800.331.8378 • www.parinc.com

Copyright © 2001, 2003 by Psychological Assessment Resources, Inc. All rights reserved. May not be reproduced in whole or in part in any form or by any means without written permission of Psychological Assessment Resources, Inc. Development of the Neuropsychological Assessment Battery was made possible and funded in part by grants 1 R43 MH58501-01 and 2 R44 MH58501-02 from the National Institute of Mental Health.

9 8 7 6 5 4 3 2 1

Reorder #RO-5089

Printed in the U.S.A.

5

Reliability and Score Differences

The statistical properties of the NAB presented in this chapter determine the confidence NAB users can have in the accuracy of obtained scores. The psychometric properties that are critical for the interpretation of test scores are reliability coefficients, standard errors of measurement (*SEM*), confidence intervals, statistical significance of the differences between scores, and frequency of score differences (i.e., base rates). This chapter focuses on a variety of factors that affect the issues of score reliability and accuracy as they relate to the quantitative interpretation of NAB performance.

RELIABILITY

The reliability of test scores refers to the accuracy, consistency, homogeneity, and temporal stability of test scores across situations or occasions (Anastasi & Urbina, 1997; Crocker & Algina, 1986). Psychometric theory and procedures allow one to estimate the magnitude of measurement error. A reliable test score will have a relatively small amount of associated measurement error. Such reliability includes both intratest and intertest consistency. As such, the reliability of a test should always be considered in the interpretation of obtained test scores and the differences between an individual's test scores on repeated testing.

Internal Consistency

First, the reliability of selected NAB tests was estimated from the item scores or subsets of item scores from a single administration. The internal consistency, or item homogeneity, was examined using alpha coefficients (Cronbach, 1951). With this method, scale consistency is determined by the interrelationships between scale items. The method also takes into account the number of items that compose the scale (Nunnally & Bernstein, 1994). Alpha coefficients produce a lower-bound estimate of the test's reliability, so this estimate is rather conservative (Allen & Yen, 1979). Many of the NAB primary subtests were excluded from internal consistency analyses for a variety of reasons including the use of unique item presentation formats, interitem dependency

issues that would artificially inflate the reliability estimate, and the use of speed of performance formats for a number of test measures. Tables 5.1 through 5.6 present alpha reliability coefficients for selected primary scores for each of the six NAB modules. For the evaluation of the internal consistency across the standardization age range, participants were categorized into the following four age groups for both Form 1 and Form 2: 18-34, 35-49, 50-69, and 70-97 years.

The range of the overall alpha coefficients is quite diverse for the six modules, when averaged for both forms across the age groups. The Form 1 and Form 2 average Screening Module alphas range from .24 for Screening Visual Discrimination (S-VIS) to .79 for Screening Digits Backward (S-DGB). Alpha coefficients for the Attention Module range from .78 for Digits Forward (DGF) to .79 for Digits Backward (DGB). Language Module alpha coefficients range from .48 for Auditory Comprehension (S-AUD) to .84 for Oral Production (OPD). Memory Module alpha coefficients range from .47 for Shape Learning Delayed Recognition (SHL-drg) to .86 for both Story Learning Phrase Unit Immediate Recall (STL-irc:phu) and Story Learning Phrase Unit Delayed Recall (STL-drc:phu). The alpha coefficients for the Spatial Module range from .65 for Map Reading (MAP) to .67 for both Visual Discrimination (VIS) and Design Construction (DES). For the Executive Functions Module, the alpha coefficients range from .45 for Judgment (JDG) to .77 for Mazes (MAZ).

It is important to recognize that there are a variety of factors that influence the magnitude of alpha coefficients. Construct heterogeneity and the range of expected scores for the particular construct of interest are of particular relevance for the evaluation of the alpha coefficients. For example, Judgment (JDG) in the Executive Functions Module represents significant item heterogeneity and would be expected to have a relatively lower alpha coefficient than Mazes (MAZ). As seen in Table 5.6, the greater diversity of item content in Judgment (JDG) yields a relatively lower alpha coefficient (i.e., overall average alpha = .45) as compared to

Table 5.1
Alpha Reliability Coefficients of Selected Screening Module Primary Scores
for the Demographically Corrected Standardization Sample by Form and Age Group

Test	Acronym	Form 1					Form 2					Forms 1 & 2 average r_{xx}^a
		Age group (years)					Age group (years)					
		18-34	35-49	50-69	70-97	Average r_{xx}^a	18-34	35-49	50-69	70-97	Average r_{xx}^a	
Screening Digits Forward	S-DGF	.77	.82	.78	.74	.78	.79	.81	.76	.75	.78	.78
Screening Digits Backward	S-DGB	.81	.83	.81	.76	.80	.77	.77	.75	.77	.77	.79
Screening Auditory Comprehension	S-AUD	.27	.30	.45	.59	.41	.71	.42	.66	.34	.55	.48
Screening Naming	S-NAM	.47	.58	.41	.52	.50	.00	.08	.28	.42	.20	.36
Screening Story Learning Immediate Recall	S-STL-irc	.62	.70	.62	.67	.65	.67	.76	.70	.76	.72	.69
Screening Story Learning Delayed Recall	S-STL-drc	.65	.74	.64	.72	.69	.69	.80	.72	.75	.74	.72
Screening Visual Discrimination	S-VIS	.27	.28	.14	.25	.24	.24	.17	.29	.27	.24	.24
Screening Design Construction	S-DES	.26	.32	.33	.32	.31	.28	.31	.38	.27	.31	.31
Screening Mazes	S-MAZ	.59	.48	.60	.61	.57	.34	.50	.60	.65	.53	.55

^a Average reliability coefficients were calculated with Fisher's z transformation.

Table 5.2
Alpha Reliability Coefficients of Selected Attention Module Primary Scores
for the Demographically Corrected Standardization Sample by Form and Age Group

Test	Acronym	Form 1					Form 2					Forms 1 & 2 average r_{xx}^a
		Age group (years)					Age group (years)					
		Average r_{xx}^a					Average r_{xx}^a					
		18-34	35-49	50-69	70-97		18-34	35-49	50-69	70-97		
Digits Forward	DGF	.77	.82	.78	.74	.78	.79	.81	.76	.75	.78	.78
Digits Backward	DGB	.81	.83	.81	.76	.80	.77	.77	.75	.77	.77	.79

^a Average reliability coefficients were calculated with Fisher's z transformation.

Table 5.3
Alpha Reliability Coefficients of Selected Language Module Primary Scores
for the Demographically Corrected Standardization Sample by Form and Age Group

Test	Acronym	Form 1					Form 2					Forms 1 & 2 average r_{xx}^a
		Age group (years)					Age group (years)					
		Average r_{xx}^a					Average r_{xx}^a					
		18-34	35-49	50-69	70-97		18-34	35-49	50-69	70-97		
Oral Production	OPD	.85	.83	.84	.82	.84	.86	.86	.86	.80	.85	.84
Auditory Comprehension	AUD	.38	.45	.49	.53	.46	.59	.40	.57	.44	.50	.48
Naming	NAM	.89	.80	.72	.70	.79	.83	.71	.70	.65	.73	.76
Bill Payment	BIL	.47	.57	.68	.79	.64	.84	.50	.74	.74	.72	.69

^a Average reliability coefficients were calculated with Fisher's z transformation.

Table 5.4
Alpha Reliability Coefficients of Selected Memory Module Primary Scores
for the Demographically Corrected Standardization Sample by Form and Age Group

Test	Acronym	Form 1					Form 2					Forms 1 & 2 average r_{xx}^a
		Age group (years)					Age group (years)					
		Average r_{xx}^a					Average r_{xx}^a					
		18-34	35-49	50-69	70-97		18-34	35-49	50-69	70-97		
Shape Learning Immediate Recognition	SHL-irg	.72	.74	.68	.56	.68	.65	.73	.76	.61	.69	.69
Shape Learning Delayed Recognition	SHL-drg	.56	.51	.48	.38	.49	.48	.46	.51	.33	.45	.47
Story Learning Phrase Unit Immediate Recall	STL-irc:phu	.88	.86	.87	.86	.87	.90	.86	.82	.85	.86	.86
Story Learning Phrase Unit Delayed Recall	STL-drc:phu	.79	.87	.88	.91	.87	.89	.85	.80	.83	.85	.86

^a Average reliability coefficients were calculated with Fisher's z transformation.

Table 5.5
Alpha Reliability Coefficients of Selected Spatial Module Primary Scores
for the Demographically Corrected Standardization Sample by Form and Age Group

Test	Acronym	Form 1					Form 2					Forms 1 & 2 average r_{xx}^a
		Age group (years)				Average r_{xx}^a	Age group (years)				Average r_{xx}^a	
		18-34	35-49	50-69	70-97		18-34	35-49	50-69	70-97		
Visual Discrimination	VIS	.71	.60	.69	.69	.67	.67	.66	.72	.62	.67	.67
Design Construction	DES	.68	.67	.65	.68	.67	.68	.71	.66	.64	.67	.67
Map Reading	MAP	.62	.68	.63	.65	.65	.45	.68	.71	.71	.65	.65

^aAverage reliability coefficients were calculated with Fisher's z transformation.

Table 5.6
Alpha Reliability Coefficients of Selected Executive Functions Module Primary Scores
for the Demographically Corrected Standardization Sample by Form and Age Group

Test	Acronym	Form 1						Form 2						Forms 1 & 2 average r_{xx}^a
		Age group (years)					Average r_{xx}^a	Age group (years)					Average r_{xx}^a	
		18-34	35-49	50-69	70-97	18-34		35-49	50-69	70-97				
Mazes	MAZ	.78	.76	.79	.80	.78	.68	.80	.79	.76	.76	.77		
Judgment	JDG	.56	.53	.45	.45	.50	.54	.48	.36	.20	.40	.45		

^aAverage reliability coefficients were calculated with Fisher's z transformation.

Mazes (overall average $\alpha = .77$) that has relatively higher item homogeneity, and thus higher item consistency.

Similarly, reduced variability of a test's scores has the effect of reducing the magnitude of alpha coefficients because, other factors aside, variability directly influences the magnitude of item intercorrelations and, thus, the average item intercorrelations. Auditory Comprehension (AUD) in the Language Module (Table 5.3), for example, has a limited range of obtained scores for nonimpaired individuals as compared to some of the other language tests. As expected, the average alpha coefficient for Auditory Comprehension (AUD; .48) is much lower relative to those of the other tests in the Language Module: .84 for Oral Production (OPD), .76 for Naming (NAM), and .69 for Bill Payment (BIL).

Test-Retest Stability

Changes in test scores from test to retest may reflect changes in an individual's cognitive ability, the effects of practice when an individual remembers the previous testing material, and/or various other situational variables related to both examiner and examinee. Consequently, the temporal stability of test scores is important because it provides information regarding the directionality and magnitude a given normative score might be expected to vary or change on retesting. The stability of NAB scores was assessed by using both Form 1 and Form 2 and sampling across a wide age range. The test-retest sample consisted of 95 individuals who were tested twice, with an average test-retest interval that exceeded 6 months ($M = 193.1$ days, $SD = 20.3$ days). The test-retest study included 45 participants who were administered Form 1 on two occasions and 50 individuals who were administered Form 2 on two occasions. The test-retest sample had the following demographic composition: 66.3% female and 33.7% male; 72.6% Caucasian, 6.3% African American, 9.5% Hispanic, and 11.6% Other race/ethnicity. The percentages of the sample by education level were 5.3%, ≤ 11 years; 21.1%, 12 years; 28.4%, 13-15 years; and 45.3%, ≥ 16 years.

The test-retest study also included an examination of stability coefficients as a function of age. For this purpose, participants were divided into two broad age groups: ages 18-59 years and ages 60-97 years. The 18-59 age group consisted of 58 participants who ranged in age from 20 to 58 years. This group had a mean age of 38.8 years ($SD = 10.7$ years) and an average education level of 14.7 years ($SD = 2.9$ years). The 60-97 age group included 37 participants who ranged in age from 60 to 97 years. These participants had a mean age of 68.9 years ($SD = 7.6$ years) and an average education level of 14.6 years ($SD = 2.4$ years). The mean scores, standard deviations, and test-retest correlations for the NAB primary, secondary, and descriptive scores are presented in Tables 5.7 through 5.12.

For the primary scores, traditional test-retest stability measures were used. Pearson correlations were first computed and then corrected for the variability of the standardization sample (see Allen & Yen, 1979; Magnusson, 1967). As with the interitem consistency estimates presented previously (i.e., alpha coefficients), there is wide variability in the test-retest stability coefficients. Average corrected stability coefficients for the primary scores across both age groups for the Screening Module range from .11 for Screening Shape Learning Immediate Recognition (S-SHL-irg) to .71 for Screening Numbers & Letters Part A Efficiency (S-N&L_A-eff). Corrected test-retest correlations for the other NAB modules averaged across all ages also demonstrated a broad range of stability coefficients. For the Attention Module, they range from .44 for Dots (DOT) to .87 for Numbers & Letters Part A Efficiency (N&L_A-eff). The Language Module coefficients range from .23 for Oral Production (OPD) to .70 for Naming (NAM). The Memory Module coefficients range from .41 for Shape Learning Delayed Recognition (SHL-drg) to .61 for Story Learning Phrase Unit Delayed Recall (STL-dr:phu). For the Spatial Module, the range is from .13 for Figure Drawing Copy (FGD-cpy) to .68 for Design Construction (DES). The Executive Functions Module coefficients range is from .43 for Judgment (JDG) to .64 for Word Generation (WGN).

As with all reliability coefficients, it is important to recognize that many factors can influence the obtained results. In particular for the test-retest study, the variability of scores can significantly affect the resulting stability coefficient. Therefore, all stability coefficients were corrected for the variability of the standardization sample. In many instances, this correction resulted in a stability coefficient that was substantially less than the obtained test-retest correlation. For example, the Screening Module results (Table 5.7) show that the test-retest correlation for Screening Numbers & Letters Part A Speed (S-N&L_A-spd) for the 18-59 age group was .66 but was corrected downward to .59. In effect, the greater-than-average variability of this group (i.e., $SD = 12.0$ at the first test, compared to $SD = 10.0$ for the standardization sample) resulted in an inflation of the obtained relationship. An example of the restriction-of-range methodology resulting in an increased stability coefficient can be seen in the Memory Module (Table 5.10) for Daily Living Memory Immediate Recall (DLM-irc) in the 60-97 age group. In this instance, the reduced variability at the first test (i.e., $SD = 7.6$ as compared to $SD = 10.0$ for the standardization sample) resulted in an upward correction of the estimated stability from .47 to .57, presumably a more accurate estimation of the relationship. Overall, the relationships between test and retest NAB scores are very good, especially given the relatively long test-retest interval, and are commensurate with, or exceed, those of most established neuropsychological measures (Nunnally & Bernstein, 1994).

Table 5.7
Stability Coefficients for Screening Module Primary, Secondary, and Descriptive Scores
for the Demographically Corrected Standardization Sample by Age Group

Test	Acronym	Age group 18-59 years ^a						Age group 60-97 years ^b						Average of age groups	
		First testing		Second testing		Reliability		First testing		Second testing		Reliability		Reliability	Corrected <i>r</i> ^d
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>r</i> ₁₂	<i>r</i> ^c	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>r</i> ₁₂	<i>r</i> ^c		
Primary score															
Screening Digits Forward	S-DGF	47.9	10.9	48.7	11.5	.75	.72	53.4	10.7	52.5	8.2	.60	.58	.65	
Screening Digits Backward	S-DGB	48.2	10.2	50.3	11.4	.73	.72	52.0	10.7	53.1	7.4	.64	.61	.67	
Screening Numbers & Letters Part A Speed	S-N&L _A -spd	52.8	12.0	49.7	12.6	.66	.59	52.7	11.2	49.8	10.4	.83	.80	.69	
Screening Numbers & Letters Part A Errors	S-N&L _A -err	48.8	11.2	52.5	7.0	.30	.27	50.0	9.7	54.4	6.1	.42	.43	.35	
Screening Numbers & Letters Part A Efficiency	S-N&L _A -eff	52.6	12.3	50.0	12.4	.69	.61	52.8	11.2	50.9	10.7	.84	.81	.71	
Screening Numbers & Letters Part B Efficiency	S-N&L _B -eff	53.4	13.4	51.8	11.8	.63	.52	52.9	9.0	52.3	10.3	.50	.54	.53	
Screening Auditory Comprehension	S-AUD	45.8	12.7	45.6	13.2	.72	.63	53.0	3.6	52.7	5.4	.08	.22	.43	
Screening Naming	S-NAM	47.9	11.3	48.5	9.8	.56	.52	51.5	7.0	52.2	7.5	.11	.15	.33	
Screening Shape Learning Immediate Recognition	S-SHL-irg	47.5	11.2	52.8	8.8	.19	.17	50.6	10.1	51.1	11.4	.05	.05	.11	
Screening Shape Learning Delayed Recognition	S-SHL-drg	50.2	10.0	51.9	9.0	.28	.28	50.5	8.8	51.2	10.3	.40	.44	.36	
Screening Story Learning Immediate Recall	S-STL-irc	45.9	10.2	50.7	8.5	.52	.51	51.9	9.0	53.0	9.1	.30	.33	.42	
Screening Story Learning Delayed Recall	S-STL-drc	47.4	10.5	52.3	10.1	.50	.48	49.2	8.7	51.3	9.7	.33	.37	.43	
Screening Visual Discrimination	S-VIS	50.6	10.5	49.6	9.4	.17	.16	49.7	9.9	52.7	8.3	.16	.16	.16	
Screening Design Construction	S-DES	49.0	10.8	50.8	10.4	.67	.64	49.8	9.5	53.4	9.9	.49	.51	.58	
Screening Mazes	S-MAZ	50.0	11.9	50.1	10.7	.69	.63	48.8	11.4	50.0	10.0	.69	.64	.63	
Screening Word Generation	S-WGN	47.4	10.9	48.8	10.7	.71	.68	48.8	9.7	52.2	10.5	.57	.58	.63	

(continued)

Table 5.7 (continued)
Stability Coefficients for Screening Module Primary, Secondary, and Descriptive Scores
for the Demographically Corrected Standardization Sample by Age Group

Test	Acronym	Age group 18-59 years ^a					Age group 60-97 years ^b					Average of age groups	
		First testing		Second testing		Reliability	First testing		Second testing		Reliability	Average of age groups	
		M	SD	M	SD		M	SD	M	SD			
						r_{12}					Corrected r^c	r_{12}	Corrected r^c
Secondary score													
Screening Orientation	S-ORN	28.9	0.4	28.9	0.5		28.7	0.6	28.7	0.6		Average % agreement ^e	89.0%
	S-DGF-sp	6.8	1.4	7.1	1.4	91.4%	7.3	1.2	6.7	1.4	86.5%		95.6%
Longest Span													
Screening Digits Backward	S-DGB-sp	5.0	1.5	5.3	1.5	96.6%	4.9	1.4	5.2	1.0	89.2%	90.3%	
Longest Span													
Screening Shape Learning	S-SHL-%rt	97.6	58.0	96.9	47.4	91.4%	91.8	62.6	75.4	60.0	81.1%	83.7%	
Percent Retention													
Screening Story Learning	S-STL-%rt	91.3	18.0	93.0	13.3	86.2%	80.2	20.6	84.6	16.5	75.7%	82.7%	
Percent Retention													
Screening Word Generation	S-WGN-psv	0.2	0.6	0.4	1.1	89.7%	0.5	0.8	0.4	0.6	70.3%	75.7%	
Perseverations													
Descriptive score													
Screening Orientation to Self	S-ORN-slf	14.0	0.1	14.0	0.1	81.0%	14.0	0.0	14.0	0.2	100.0%	100.0%	
Screening Orientation to Time	S-ORN-tim	10.0	0.3	9.9	0.3	100.0%	9.8	0.6	9.9	0.2	91.9%	94.3%	
Screening Orientation to Place	S-ORN-plc	3.9	0.2	3.9	0.4	96.6%	4.0	0.0	3.8	0.5	100.0%	100.0%	
Screening Orientation to Situation	S-ORN-sit	1.0	0.0	1.0	0.0	100.0%	1.0	0.2	1.0	0.0	100.0%	100.0%	
Screening Auditory Comprehension Colors	S-AUD-col	13.0	0.0	13.0	0.3	100.0%	13.0	0.0	13.0	0.0	100.0%	100.0%	
Screening Auditory Comprehension Shapes	S-AUD-shp	21.8	0.6	21.8	0.9	96.6%	22.0	0.2	21.9	0.3	100.0%	98.3%	
Screening Auditory Comprehension Colors/Shapes/Numbers	S-AUD-csn	20.9	0.4	20.9	0.6	98.3%	20.9	0.4	20.9	0.3	100.0%	99.2%	
Screening Naming Percent Correct After Semantic Cuing	S-NAM-sem%	20.5	31.3	42.5	49.4	74.1%	34.5	45.5	28.7	41.9	56.8%	65.5%	
Screening Naming Percent Correct After Phonemic Cuing	S-NAM-pho%	83.3	38.9	75.8	39.4	86.2%	70.0	35.0	67.9	46.4	73.0%	79.6%	

^a $N = 58$. ^b $N = 37$. ^cCorrelations were corrected for restriction of range using the initial (i.e., first) test (Nunnally & Bernstein, 1994). ^dArithmetic average of each corrected reliability for all age groups. ^eSecondary and descriptive scores are reported as percentage agreement.

Table 5.8
Stability Coefficients for Attention Module Primary, Secondary, and Descriptive Scores
for the Demographically Corrected Standardization Sample by Age Group

Test	Acronym	Age group 18-59 years ^a						Age group 60-97 years ^b						Average of age groups	
		First testing			Second testing			First testing			Second testing			Reliability	
		Reliability			Reliability			Reliability			Reliability			Reliability	
		M	SD	r_{12}	M	SD	Corrected r^c	M	SD	r_{12}	M	SD	Corrected r^c	Corrected r^d	
Primary score															
Digits Forward	DGF	47.9	10.9	48.7	11.5	.75	.72	53.4	10.7	52.5	8.2	.60	.58	.65	
Digits Backward	DGB	48.2	10.2	50.3	11.4	.73	.72	52.0	10.7	53.1	7.4	.64	.61	.67	
Dots	DOT	48.0	10.6	51.5	10.8	.52	.49	48.8	10.9	50.7	8.4	.41	.38	.44	
Numbers & Letters Part A Speed	N&L _A -spd	52.1	12.5	52.3	12.4	.92	.88	49.4	8.5	49.5	9.3	.81	.85	.86	
Numbers & Letters Part A Errors	N&L _A -err	48.8	11.6	51.6	9.6	.66	.60	53.6	8.5	52.8	9.9	.47	.53	.56	
Numbers & Letters Part A Efficiency	N&L _A -eff	51.8	12.3	52.3	12.5	.89	.85	50.1	8.2	49.5	9.5	.85	.89	.87	
Numbers & Letters Part B Efficiency	N&L _B -eff	48.3	10.5	52.0	10.8	.56	.53	51.5	8.3	51.2	8.7	.55	.62	.58	
Numbers & Letters Part C Efficiency	N&L _C -eff	48.7	11.5	53.0	10.6	.57	.51	49.8	10.4	52.7	10.0	.66	.64	.58	
Numbers & Letters Part D Efficiency	N&L _D -eff	49.8	12.0	52.0	11.8	.81	.75	50.9	9.6	52.2	7.6	.53	.54	.65	
Numbers & Letters Part D Disruption	N&L _D -dis	51.0	9.9	52.2	10.6	.38	.39	50.6	10.8	52.9	8.1	.59	.56	.47	
Driving Scenes	DRV	45.1	9.7	49.2	9.4	.71	.72	47.2	10.8	48.8	8.1	.59	.56	.64	
Secondary score															
Orientation	ORN	28.9	0.4	28.9	0.5	91.4%	91.4%	28.7	0.6	28.7	0.6	86.5%	86.5%	89.0%	
Digits Forward Longest Span	DGF-spn	6.8	1.4	7.1	1.4	96.6%	96.6%	7.3	1.2	6.7	1.4	94.6%	94.6%	95.6%	
Digits Backward Longest Span	DGB-spn	5.0	1.5	5.3	1.5	91.4%	91.4%	4.9	1.4	5.2	1.0	89.2%	89.2%	90.3%	
Descriptive score															
Orientation to Self	ORN-slf	14.0	0.1	14.0	0.1	100.0%	100.0%	14.0	0.0	14.0	0.2	100.0%	100.0%	100.0%	
Orientation to Time	ORN-tim	10.0	0.3	9.9	0.3	96.6%	96.6%	9.8	0.6	9.9	0.2	91.9%	91.9%	94.3%	
Orientation to Place	ORN-plc	3.9	0.2	3.9	0.4	100.0%	100.0%	4.0	0.0	3.8	0.5	100.0%	100.0%	100.0%	
Orientation to Situation	ORN-sit	1.0	0.0	1.0	0.0	100.0%	100.0%	1.0	0.2	1.0	0.0	100.0%	100.0%	100.0%	

^a $N = 58$. ^b $N = 37$. ^cCorrelations were corrected for restriction of range using the initial (i.e., first) test (Nunnally & Bernstein, 1994). ^dArithmetic average of each corrected reliability for all age groups. ^eSecondary and descriptive scores are reported as percentage agreement.

Table 5.9

Stability Coefficients for Language Module Primary, Secondary, and Descriptive Scores for the Demographically Corrected Standardization Sample by Age Group

Test	Acronym	Age group 18-59 years ^a						Age group 60-97 years ^b						Average of age groups	
		First testing			Second testing			First testing			Second testing			Reliability	
		M	SD	Corrected r^c	M	SD	Corrected r^c	M	SD	Corrected r^c	M	SD	Corrected r^c	Corrected r^d	
															Reliability
Primary score															
Oral Production	OPD	42.3	10.1	46.3	8.4	.16	.16	44.4	7.6	48.5	8.6	.23	.30	.23	
Auditory Comprehension	AUD	47.9	9.9	48.9	10.0	.25	.26	50.9	8.9	54.5	5.4	.25	.27	.26	
Naming	NAM	47.2	11.9	46.4	11.4	.75	.69	52.0	9.3	52.3	9.0	.69	.72	.70	
Writing	WRT	48.7	11.7	48.8	11.8	.32	.28	50.1	9.3	52.1	8.2	.19	.20	.24	
Bill Payment	BIL	48.2	10.9	49.4	10.7	.55	.51	52.1	7.9	51.2	9.8	.25	.32	.41	
Secondary score															
Reading Comprehension	RCN	13.0	0.2	12.9	0.4	98.3%	98.3%	12.9	0.3	13.0	0.0	100.0%	100.0%	99.2%	
Writing Legibility	WRT-leg	2.0	0.1	1.9	0.2	100.0%	100.0%	2.0	0.2	1.9	0.3	100.0%	100.0%	100.0%	
Writing Spelling	WRT-spl	2.8	0.5	2.6	0.5	96.6%	96.6%	2.7	0.5	2.7	0.5	100.0%	100.0%	98.3%	
Writing Syntax	WRT-syn	2.9	0.3	2.9	0.2	100.0%	100.0%	2.9	0.3	3.0	0.2	100.0%	100.0%	100.0%	
Writing Conveyance	WRT-cnv	2.4	0.7	2.5	0.7	87.9%	87.9%	2.4	0.6	2.6	0.6	89.2%	89.2%	88.6%	
Descriptive score															
Auditory Comprehension Colors	AUD-col	13.0	0.0	13.0	0.3	100.0%	100.0%	13.0	0.0	13.0	0.0	100.0%	100.0%	100.0%	
Auditory Comprehension Shapes	AUD-shp	21.8	0.6	21.8	0.9	96.6%	96.6%	22.0	0.2	21.9	0.3	100.0%	100.0%	98.3%	
Auditory Comprehension Colors/ Shapes/Numbers	AUD-csn	20.9	0.4	20.9	0.6	98.3%	98.3%	20.9	0.4	20.9	0.3	100.0%	100.0%	99.2%	
Auditory Comprehension Pointing	AUD-pnt	6.0	0.0	6.0	0.0	100.0%	100.0%	6.0	0.0	6.0	0.0	100.0%	100.0%	100.0%	
Auditory Comprehension Yes/No	AUD-y/n	9.6	0.9	9.7	0.9	94.8%	94.8%	9.5	1.0	9.7	0.7	97.3%	97.3%	96.1%	
Auditory Comprehension Paper Folding	AUD-flid	16.2	1.1	16.3	1.3	94.8%	94.8%	15.9	1.5	16.2	1.0	86.5%	86.5%	90.7%	
Naming Percent Correct After Semantic Cuing	NAM-sem%	19.8	32.9	37.0	35.9	77.6%	77.6%	30.1	36.3	19.9	33.9	81.1%	81.1%	79.4%	
Naming Percent Correct After Phonemic Cuing	NAM-pho%	60.6	37.8	61.0	40.5	77.6%	77.6%	69.7	42.0	64.6	36.1	73.0%	73.0%	75.3%	
Reading Comprehension Words	RCN-wrd	6.0	0.0	6.0	0.1	98.3%	98.3%	6.0	0.0	6.0	0.0	100.0%	100.0%	99.2%	
Reading Comprehension Sentences	RCN-sen	7.0	0.2	6.9	0.3	98.3%	98.3%	6.9	0.3	7.0	0.0	97.3%	97.3%	97.8%	

^a $N = 58$. ^b $N = 37$. ^cCorrelations were corrected for restriction of range using the initial (i.e., first) test (Nunnally & Bernstein, 1994). ^dArithmetic average of each corrected reliability for all age groups. ^eSecondary and descriptive scores are reported as percentage agreement.

Table 5.10
Stability Coefficients for Memory Module Primary, Secondary, and Descriptive Scores
for the Demographically Corrected Standardization Sample by Age Group

Test	Acronym	Age group 18-59 years ^a						Age group 60-97 years ^b						Average of age groups	
		First testing			Second testing			First testing			Second testing			Reliability	
		M	SD	r ₁₂	M	SD	r ₁₂	M	SD	r ₁₂	M	SD	r ₁₂	r ^c	Reliability
Primary score															
List Learning List A Immediate Recall	LLA-irc	48.2	9.9	50.6	9.7	.48	.49	48.7	7.0	50.4	7.8	.34	.46	.47	
List Learning List B Immediate Recall	LLB-irc	48.2	9.8	46.8	10.6	.36	.36	50.6	10.8	49.9	8.5	.62	.59	.48	
List Learning List A Short Delayed Recall	LLA-sd:drc	48.1	10.1	49.3	10.6	.54	.54	50.6	7.5	52.3	7.0	.27	.35	.44	
List Learning List A Long Delayed Recall	LLA-ls:drc	48.1	10.4	50.7	9.1	.43	.41	50.2	7.6	52.7	7.4	.64	.74	.58	
Shape Learning Immediate Recognition	SHL-irg	50.7	8.9	53.9	12.2	.35	.38	52.3	12.6	56.4	10.0	.55	.47	.43	
Shape Learning Delayed Recognition	SHL-drg	50.1	9.8	52.2	11.6	.34	.34	50.5	11.3	52.3	11.0	.52	.47	.41	
Story Learning Phrase Unit Immediate Recall	STL-irc:phu	48.1	10.2	52.1	9.4	.70	.69	50.1	9.3	49.7	6.9	.33	.36	.52	
Story Learning Phrase Unit Delayed Recall	STL-drc:phu	49.1	8.7	52.9	9.3	.67	.72	52.4	9.3	52.5	8.7	.47	.50	.61	
Daily Living Memory Immediate Recall	DLM-irc	48.4	10.9	54.4	9.6	.50	.46	53.0	7.6	56.1	8.7	.47	.57	.52	
Daily Living Memory Delayed Recall	DLM-drc	47.7	11.6	48.7	10.2	.51	.45	52.6	9.0	55.4	8.2	.45	.49	.47	
Secondary score															
List Learning List A Trial 1 Immediate Recall	LLA1-irc	6.2	1.7	6.6	1.8	.966%	.966%	5.3	1.4	5.4	1.7	86.5%	86.5%	91.6%	
List Learning List A Trial 2 Immediate Recall	LLA2-irc	9.1	2.1	9.6	2.2	.931%	.931%	8.0	1.8	8.1	1.9	86.5%	86.5%	89.8%	
List Learning List A Trial 3 Immediate Recall	LLA3-irc	10.3	1.8	10.4	2.0	.948%	.948%	9.1	1.9	9.5	1.8	91.9%	91.9%	93.4%	
Average agreement ^e															
91.6%															
89.8%															
93.4%															
(continued)															

(continued)

Table 5.10 (continued)
Stability Coefficients for Memory Module Primary, Secondary, and Descriptive Scores
for the Demographically Corrected Standardization Sample by Age Group

Test	Acronym	Age group 18-59 years ^a						Age group 60-97 years ^b						Average of age groups	
		First testing			Second testing			First testing			Second testing			Reliability	
		M	SD	r ₁₂	M	SD	r ^c	M	SD	r ₁₂	M	SD	r ^c	Corrected	r ^d
List Learning List A Percent Retention	LLA-%rt	86.1	19.6	90.0	24.6	93.1%	83.2	19.3	86.8	16.8	78.4%	85.8%			
List Learning List A Long Delayed Forced-Choice Recognition	LLA-dfc	11.4	1.1	11.5	1.0	96.6%	11.1	1.1	11.4	0.9	86.5%	91.6%			
List Learning List A Long Delayed Forced-Choice Recognition False Alarms	LLA-fa	2.0	3.2	1.5	2.5	93.1%	2.7	2.4	2.8	2.9	83.8%	88.5%			
List Learning List A Discriminability Index	LLA-dis	9.4	3.5	10.0	2.8	91.4%	8.4	2.9	8.6	3.1	97.3%	94.4%			
List Learning List A Recall vs. Recognition	LLA-rvr	78.3	20.8	80.6	21.3	91.4%	67.5	17.9	72.2	17.5	89.2%	90.3%			
Shape Learning Trial 1 Immediate Recognition	SHL1-irg	5.4	1.3	6.0	1.7	100.0%	4.2	1.6	5.1	1.5	86.5%	93.3%			
Shape Learning Trial 2 Immediate Recognition	SHL2-irg	6.5	1.4	6.8	2.0	93.1%	5.4	2.0	6.1	1.5	78.4%	85.8%			
Shape Learning Trial 3 Immediate Recognition	SHL3-irg	6.9	1.5	7.2	1.6	96.6%	6.2	1.8	6.4	1.7	89.2%	92.9%			
Shape Learning Percent Retention	SHL-%rt	98.8	21.3	99.1	18.8	87.9%	89.5	29.7	91.2	21.1	56.8%	72.4%			
Shape Learning Delayed Forced-Choice Recognition	SHL-dfc	8.1	1.2	8.2	1.0	91.4%	7.4	1.3	7.9	1.0	86.5%	89.0%			
Shape Learning Delayed Forced-Choice Recognition False Alarms	SHL-fa	0.1	0.4	0.1	0.4	96.6%	0.6	1.1	0.5	1.0	81.1%	88.9%			
Shape Learning Discriminability Index	SHL-dis	7.9	1.2	8.1	1.1	98.3%	6.8	1.6	7.5	1.4	91.9%	95.1%			
Story Learning Trial 1 Phrase Unit	STL1-irc:phu	24.4	7.6	27.1	6.0	81.0%	23.8	6.6	23.6	5.4	86.5%	83.8%			
Story Learning Trial 2 Phrase Unit	STL2-irc:phu	32.9	5.8	34.7	4.8	94.8%	31.2	7.5	31.5	5.8	86.5%	90.7%			
Story Learning Thematic Unit Immediate Recall	STL-irc:thu	15.7	3.4	16.5	2.6	96.6%	14.6	3.7	15.3	3.0	89.2%	92.9%			
(continued)															

(continued)

Table 5.10 (continued)
Stability Coefficients for Memory Module Primary, Secondary and Descriptive Scores
for the Demographically Corrected Standardization Sample by Age Group

Test	Acronym	Age group 18-59 years ^a						Age group 60-97 years ^b						Average of age groups				
		First testing			Second testing			Reliability			First testing			Second testing			Reliability	
								Corrected									Corrected	
		M	SD		M	SD		r ₁₂	r ^c		M	SD		M	SD		r ₁₂	r ^c
Story Learning Trial 1 Thematic Unit	STL1-irc:thu	6.9	2.2	7.4	1.6		87.9%		6.2	2.1	6.8	1.5		94.6%		91.3%		
Story Learning Trial 2 Thematic Unit	STL2-irc:thu	8.8	1.5	9.1	1.2		100.0%		8.4	2.0	8.4	1.7		97.3%		98.7%		
Story Learning Thematic Unit Delayed Recall	STL-drc:thu	8.6	1.7	8.9	1.4		98.3%		8.0	2.1	8.3	1.8		94.6%		96.5%		
Story Learning Phrase Unit Percent Retention	STL-%rt	92.8	9.7	95.2	7.0		93.1%		94.9	17.3	91.8	10.9		83.8%		88.5%		
Daily Living Memory Retention	DLM-rt	93.2	12.1	91.8	12.9		89.7%		89.8	11.5	90.7	11.5		89.2%		89.5%		
Daily Living Memory Delayed Recognition	DLM-drg	9.2	1.1	9.4	0.9		86.2%		8.8	1.2	9.2	0.9		89.2%		87.7%		
Daily Living Memory Recall vs. Recognition	DLM-rvr	163.6	22.2	161.7	20.0		86.2%		165.9	18.3	161.8	17.7		97.3%		91.8%		
Medication Instructions Immediate Recall	MED-irc	24.2	2.5	24.9	2.5		93.1%		24.3	2.0	24.2	2.2		91.9%		92.5%		
Medication Instructions Delayed Recall	MED-drc	8.4	1.1	8.3	1.3		91.4%		8.2	0.9	8.3	0.9		94.6%		93.0%		
Medication Instructions Delayed Recognition	MED-drg	1.5	0.7	1.7	0.5		87.9%		1.4	0.7	1.6	0.5		83.8%		85.9%		
Name/Address/Phone Immediate Recall	NAP-irc	19.7	3.9	21.5	2.5		91.4%		19.1	2.6	20.3	2.9		94.6%		93.0%		
Name/Address/Phone Delayed Recall	NAP-drc	6.8	1.9	7.1	1.6		91.4%		6.3	1.8	6.6	1.7		100.0%		95.7%		
Name/Address/Phone Delayed Recognition	NAP-drg	7.7	0.7	7.8	0.7		94.8%		7.4	1.0	7.6	0.6		91.9%		93.4%		
Descriptive score																		
List Learning Semantic Clusters	LL-sem	16.6	9.0	18.0	9.6		89.7%		13.5	6.1	14.0	7.2		91.9%		90.8%		
List Learning Perseverations	LL-psv	2.3	2.5	2.0	2.2		63.8%		1.9	1.7	2.6	3.0		73.0%		68.4%		
List Learning Intrusions	LL-int	1.3	1.8	1.3	2.0		82.8%		1.1	1.4	1.1	1.1		83.8%		83.3%		

^a*N* = 58. ^b*N* = 37. ^cCorrelations were corrected for restriction of range using the initial (i.e., first) test (Nunnally & Bernstein, 1994). ^dArithmetic average of each corrected reliability for all age groups. ^eSecondary and descriptive scores are reported as percentage agreement.

Table 5.11
Stability Coefficients for Spatial Module Primary and Secondary Scores
for the Demographically Corrected Standardization Sample by Age Group

Test	Acronym	Age group 18-59 years ^a						Age group 60-97 years ^b						Average of age groups	
		First testing			Second testing			First testing			Second testing			Reliability	
		M	SD	r ₁₂	M	SD	r _c	M	SD	r ₁₂	M	SD	r _c	r ₁₂	r _c
Primary score															
Visual Discrimination	VIS	49.3	10.6	48.3	11.1	.32	.30	52.3	10.0	.45	.45	.38			
Design Construction	DES	47.4	11.1	50.4	11.3	.66	.62	49.0	12.6	.81	.74	.68			
Figure Drawing Copy	FGD-cpy	51.7	9.2	42.4	12.4	.17	.19	51.1	10.8	.08	.08	.13			
Figure Drawing Copy Organization	FGD-cpy:org	49.0	10.0	50.0	9.8	.44	.44	50.5	9.2	.23	.25	.35			
Figure Drawing Immediate Recall	FGD-irc	48.8	10.1	46.4	12.4	.45	.44	50.0	9.6	.61	.63	.54			
Map Reading	MAP	48.5	11.7	46.9	10.6	.55	.49	48.5	10.2	.60	.59	.54			
Secondary score															
Figure Drawing Copy Fragmentation	FGD-cpy:frg	4.2	1.0	4.2	1.2	87.9%	87.9%	4.4	0.9	89.2%	89.2%	88.6%			
Figure Drawing Copy Planning	FGD-cpy:pln	3.2	1.0	3.3	0.8	84.5%	84.5%	3.2	0.8	86.5%	86.5%	85.5%			
Figure Drawing Immediate Recall Organization	FGD-irc:org	6.9	2.1	7.3	2.0	81.0%	81.0%	6.5	1.7	73.0%	73.0%	77.0%			
Figure Drawing Immediate Recall Fragmentation	FGD-irc:frg	4.0	1.2	4.3	1.1	89.7%	89.7%	3.8	1.0	78.4%	78.4%	84.1%			
Figure Drawing Immediate Recall Planning	FGD-irc:pln	2.9	1.2	3.1	1.1	89.7%	89.7%	2.7	0.8	91.9%	91.9%	90.8%			
Figure Drawing Percent Retention	FGD-%rt	78.9	16.5	82.1	21.8	81.0%	81.0%	73.4	17.4	64.9%	64.9%	73.0%			

^a*N* = 58. ^b*N* = 37. ^cCorrelations were corrected for restriction of range using the initial (i.e., first) test (Nunnally & Bernstein, 1994). ^dArithmetic average of each corrected reliability for all age groups. ^eSecondary scores are reported as percentage agreement.

Table 5.12
Stability Coefficients for Executive Functions Module Primary and Secondary Scores
for the Demographically Corrected Standardization Sample by Age Group

Test	Acronym	Age group 18-59 years ^a						Age group 60-97 years ^b						Average of age groups	
		First testing			Second testing			First testing			Second testing			Reliability	
		M	SD	r ₁₂	M	SD	r ₁₂	M	SD	r ₁₂	M	SD	r ₁₂	Reliability	
														Corrected r ^c	Corrected r ^c
Primary score													Corrected r ^d		
Mazes	MAZ	50.3	13.2	.69	52.2	13.0	.59	49.0	11.5	.61	50.5	8.7	.55	.57	
Judgment	JDG	47.7	10.0	.49	43.0	9.8	.49	47.5	8.7	.32	46.9	10.4	.37	.43	
Categories	CAT	47.3	10.7	.49	50.4	11.5	.47	48.3	8.3	.53	54.0	9.8	.60	.54	
Word Generation	WGN	49.7	11.7	.65	50.0	10.8	.59	49.8	8.4	.63	52.5	7.6	.70	.64	
Secondary score				% agreement ^e			% agreement ^e			% agreement ^e				Average % agreement ^e	
	Word Generation Perseverations	WGN-psv	1.0	1.4	81.0%	1.0	1.5	81.0%	1.2	1.3	83.8%	1.5	1.6	82.4%	82.4%

^a*N* = 58. ^b*N* = 37. ^cCorrelations were corrected for restriction of range using the initial (i.e., first) test (Nunnally & Bernstein, 1994). ^dArithmetic average of each corrected reliability for all age groups. ^eSecondary scores are reported as percentage agreement.

Additionally, an examination of the mean level of performance from test to retest shows relatively small changes (see Tables 5.7 through 5.12). As was expected, participants tended to perform slightly better on the second testing, thus demonstrating a small practice effect. Although most changes were relatively small, there was a trend for greater practice effects in the Memory and Language Modules. Overall, however, the remarkable finding is the similarity in level of performance from test to retest. One factor that contributes to the small practice effects is the relatively long test–retest interval, which exceeded an average of 6 months. The 6-month retest interval was chosen because it was thought to correspond to the most frequent retest interval encountered in clinical practice. Test–retest intervals that are substantially shorter in duration may be expected to produce greater practice effects. Users of the NAB can have high confidence that practice effects will have a minimal effect when used in reevaluation situations, especially with test–retest intervals of several months or longer.

In order to examine the temporal stability of the secondary and descriptive scores, a percentage agreement coefficient was calculated. The results are also shown in Tables 5.7 through 5.12. Percentage agreement coefficients are sometimes referred to as “decision consistency reliability” and are an alternate way of evaluating the stability of test–retest measures. The primary reason for using the percentage agreement methodology for secondary and descriptive scores was the nature of many of the variables (i.e., the use of process measures that include many difference scores or ratios and many score distributions that were extremely skewed or had very limited variability). For purposes of this stability study, the test and retest scores were divided into three categories or classification ranges based on standard deviation units that are commonly used to denote the neuropsychological functioning ranges of moderate-to-severe impairment and worse (≤ -2.0 *SD*), moderate impairment to above average performance (-1.9 *SD* to $+1.9$ *SD*), and superior performance (≥ 2.0 *SD*). The test and retest ranges were then evaluated to determine the percentage of agreement of each category from test to retest. The resulting percentage agreement coefficient reflects the percentage of temporal consistency for the classification range as opposed to the stability of the individual’s obtained score.

As the data in Tables 5.7 through 5.12 show, the percentage agreement coefficients are generally quite high. For the secondary scores, the median percentage agreement coefficients for correct classifications (i.e., based on the average correct classification percentage for all ages) was 86.4% (Screening Module), 90.3% (Attention Module), 99.2% (Language Module), 91.6% (Memory Module), 84.8% (Spatial Module), and 82.4% (Executive Functions Module).

Similarly, the median percentage agreement coefficients for the descriptive scores were 99.2% (Screening Module), 100.0% (Attention Module), 98.3% (Language Module), and 83.3% (Memory Module). These data demonstrate that the temporal consistency of range classification is excellent and that the NAB user can have high confidence in obtaining similar test–retest classification ranges.

Tables 5.13 through 5.17 present the stability coefficients for the NAB Screening Domain, Total Screening Index, module index, and Total NAB Index scores. Tables 5.13 and 5.14 show the Screening Domain stability coefficients for the demographically corrected standardization sample and the age-based, U.S. Census-matched standardization sample, respectively. As expected, the stability coefficients for the Screening Domain scores and the Total Screening Index score are more stable than the individual test scores that contribute to the composite score. The average test–retest coefficients for the Screening Domain scores are quite comparable for both standardization samples in that both samples have median test–retest correlations of .60 averaged across the age groups. Similarly, the Total Screening Index score for the demographically corrected standardization sample and the U.S. Census-matched standardization sample has test–retest correlations of .74 and .75 averaged across age groups, respectively.

The NAB stability coefficients for the module indexes are presented in Tables 5.15 and 5.16 for the demographically corrected standardization and U.S. Census-matched standardization samples, respectively. Similar to the Screening Domain scores and the Total Screening Index score, the module indexes generally have higher correlations than their component primary test scores. The mean test–retest coefficients averaged across age groups for the module index scores are comparable for the demographically corrected standardization (.69) and U.S. Census-matched standardization samples (.75). The average stability coefficients for the Total NAB Index for the demographically corrected standardization sample (.82) and the U.S. Census-matched sample (.86) are quite high.

The data in Tables 5.13 and 5.14 generally show that the Screening Domain scores have relatively small practice effects from test to retest. Most practice effects are less than one half of a standard deviation for both the demographically corrected standardization sample and U.S. Census-matched sample. The notable exception is the Screening Memory Domain score for the younger age group (18-59 years). As demonstrated by the increase in means from test to retest, there is about a one standard deviation increase in *T*-score points (i.e., practice effect). In contrast, the Screening Memory Domain score practice effect is much smaller for the older age group (60-97 years), which had

Table 5.13
Stability Coefficients for the Screening Module Domain and
Total Screening Index Scores for the Demographically Corrected Standardization Sample by Age Group

Domain/Index score	Acronym	Age group 18-59 years ^a						Age group 60-97 years ^b						Average of age groups	
		First testing		Second testing		Reliability	First testing		Second testing		Reliability	Reliability	Corrected r^d		
		M	SD	M	SD		M	SD	M	SD				M	SD
						Corrected r^c					Corrected r^c	Corrected r^c	Corrected r^c		
Screening Attention Domain	S-ATT	101.1	17.5	100.5	17.3	.73	.67	106.2	13.3	104.9	12.4	.73	.77	.72	
Screening Language Domain	S-LAN	93.6	17.5	93.5	16.9	.68	.62	101.4	10.3	102.2	10.3	.60	.73	.68	
Screening Memory Domain	S-MEM	95.1	14.9	103.9	14.2	.57	.57	100.9	12.2	103.3	13.4	.40	.47	.52	
Screening Spatial Domain	S-SPT	99.7	15.5	100.2	14.2	.60	.58	99.6	15.7	105.9	15.1	.54	.53	.55	
Screening Executive Functions Domain	S-EXE	97.8	18.4	99.1	16.1	.74	.67	97.8	15.4	102.2	15.1	.55	.54	.60	
Total Screening Index	S-NAB	96.3	16.7	99.1	16.1	.79	.75	101.6	11.9	106.0	13.7	.64	.73	.74	

^a $N = 58$. ^b $N = 37$. ^cCorrelations were corrected for restriction of range using the initial (i.e., first) test (Nunnally & Bernstein, 1994). ^dArithmetic average of each corrected reliability for all age groups.

Table 5.14
Stability Coefficients for the Screening Module Domain and
Total Screening Index Scores for the Age-Based, U.S. Census-Matched Sample by Age Group

Domain/Index score	Acronym	Age group 18-59 years ^a					Age group 60-97 years ^b					Average of age groups	
		First testing		Second testing		Reliability	First testing		Second testing		Reliability	Reliability	
		M	SD	M	SD		M	SD	M	SD			
						Corrected r^c					Corrected r^c		Corrected r^d
Screening Attention Domain	S-ATT	103.2	16.2	102.5	16.3	.71	108.5	13.3	106.9	12.4	.75	.79	.74
Screening Language Domain	S-LAN	94.4	13.2	94.7	12.4	.46	102.3	11.5	103.1	11.0	.57	.68	.59
Screening Memory Domain	S-MEM	97.6	15.7	106.1	14.3	.62	103.0	12.6	105.2	12.0	.43	.49	.54
Screening Spatial Domain	S-SPT	100.9	15.2	101.0	14.3	.64	101.0	16.4	107.1	14.9	.59	.56	.60
Screening Executive Functions Domain	S-EXE	99.4	17.5	100.8	15.4	.77	99.8	15.9	103.7	15.2	.62	.60	.66
Total Screening Index	S-NAB	98.9	15.6	101.6	15.3	.78	104.4	13.1	108.1	13.2	.69	.74	.75

^a $N = 58$. ^b $N = 37$. ^cCorrelations were corrected for restriction of range using the initial (i.e., first) test (Nunnally & Bernstein, 1994). ^dArithmetic average of each corrected reliability for all age groups.

Table 5.15
Stability Coefficients for the Module and Total NAB Index Scores for the
Demographically Corrected Standardization Sample by Age Group

Index score	Acronym	Age group 18-59 years ^a						Age group 60-97 years ^b						Average of age groups	
		First testing			Second testing			First testing			Second testing			Reliability	
														Corrected	
		M	SD		M	SD		M	SD		M	SD		r_{12}	r^c
Attention Index	ATT	96.2	16.6	102.7	16.0	.88	101.0	13.1	102.9	11.1	.81	.84			.85
Language Index	LAN	91.7	14.0	94.7	14.3	.55	98.4	12.0	103.1	11.7	.44	.52			.55
Memory Index	MEM	97.0	14.9	103.3	15.2	.74	101.7	11.5	106.1	12.6	.50	.60			.67
Spatial Index	SPT	98.2	16.7	93.7	16.3	.66	100.2	14.3	96.2	15.8	.78	.80			.71
Executive Functions Index	EXE	97.4	16.2	97.5	18.2	.68	96.7	13.3	101.7	13.4	.69	.73			.69
Total NAB Index	T-NAB	95.1	15.1	98.0	17.0	.80	99.3	13.2	102.4	12.1	.82	.85			.82

^a $N = 58$. ^b $N = 37$. ^cCorrelations were corrected for restriction of range using the initial (i.e., first) test (Nunnally & Bernstein, 1994). ^dArithmetic average of each corrected reliability for all age groups.

Table 5.16
Stability Coefficients for the Module and Total NAB Index Scores
for the Age-Based, U.S. Census-Matched Sample by Age Group

Index score	Acronym	Age group 18-59 years ^a						Age group 60-97 years ^b						Average of age groups	
		First testing			Second testing			First testing			Second testing			Reliability	
														Corrected	
														r^c	
		M	SD		M	SD		M	SD		M	SD		r_{12}	r^c
Attention Index	ATT	98.4	15.9	104.6	15.4	.89	103.5	12.4	105.3	10.7	.81	.86	.87		
Language Index	LAN	94.3	12.6	97.4	14.3	.56	100.6	12.3	104.9	12.3	.56	.64	.63		
Memory Index	MEM	100.2	15.4	105.6	16.1	.82	105.1	11.0	109.2	12.4	.54	.66	.74		
Spatial Index	SPT	99.8	16.7	95.7	16.0	.71	101.9	14.2	98.6	15.0	.81	.83	.75		
Executive Functions Index	EXE	99.6	16.2	99.7	18.2	.76	100.4	12.4	104.8	12.8	.70	.76	.75		
Total NAB Index	T-NAB	98.0	15.2	100.8	17.1	.85	102.7	12.8	105.4	11.8	.83	.87	.86		

^a $N = 58$. ^b $N = 37$. ^cCorrelations were corrected for restriction of range using the initial (i.e., first) test (Nunnally & Bernstein, 1994). ^dArithmetic average of each corrected reliability for all age groups.

Table 5.17
Stability Coefficients of the Screening Domain, Total Screening Index, Module Index, and
Total NAB Index Scores for the Demographically Corrected Standardization Sample by Form

Domain/Index	Acronym	Form 1					Form 2					Average of forms		
		First testing		Second testing		Reliability	First testing		Second testing		Reliability	Average of forms		
		M	SD	M	SD		M	SD	M	SD				
						r^a					r^b			
Screening Domain score														
Screening Attention Domain	S-ATT	101.2	15.9	101.8	16.4	.64	.61	104.8	16.3	102.5	15.1	.83	.81	.71
Screening Language Domain	S-LAN	102.1	18.4	103.9	16.8	.56	.48	91.6	10.1	90.6	10.2	.84	.92	.70
Screening Memory Domain	S-MEM	96.9	14.1	102.3	13.2	.63	.65	97.7	14.3	104.8	14.3	.39	.40	.53
Screening Spatial Domain	S-SPT	103.2	15.6	103.3	17.1	.68	.66	96.5	14.8	101.6	12.5	.43	.43	.55
Screening Executive Functions Domain	S-EXE	96.7	19.6	98.3	17.4	.67	.56	98.7	14.7	102.0	14.0	.68	.69	.62
Total Screening Index	S-NAB	100.3	17.0	103.4	17.2	.71	.66	96.6	13.3	100.3	13.8	.82	.85	.75
Index score														
Attention Index	ATT	97.9	15.4	102.9	14.0	.87	.86	98.2	15.6	102.8	14.5	.84	.83	.85
Language Index	LAN	95.0	14.0	94.5	14.2	.52	.55	93.8	13.3	101.1	13.0	.63	.67	.61
Memory Index	MEM	97.5	14.2	104.0	15.7	.63	.65	100.0	13.5	104.8	12.9	.73	.76	.71
Spatial Index	SPT	98.6	14.5	92.6	14.8	.62	.63	99.4	16.9	96.5	17.0	.76	.72	.68
Executive Functions Index	EXE	95.1	16.3	96.9	17.4	.59	.56	98.9	13.7	101.2	15.6	.76	.79	.67
Total NAB Index	T-NAB	96.5	14.9	97.8	15.8	.74	.74	97.0	14.2	101.4	14.8	.88	.89	.81

^aCorrelations were corrected for restriction of range using the initial (i.e., first) test (Nunnally & Bernstein, 1994). ^bArithmetic average of each corrected reliability for all age groups.

an average increase of only 2 *T*-score points at retest. These results, and in particular the age trends, should be taken into consideration in the interpretation of NAB retest memory scores. With the exception of the Screening Memory Domain score, the NAB Screening Domain scores and the Total Screening Index demonstrate relatively small practice effects and, in most cases, the mean increase in retest scores is well within the expected standard error of measurement.

An evaluation of the test–retest mean scores in Tables 5.15 and 5.16 reveal a similar pattern of practice effects for the module index scores and the Total NAB Index score. The practice effects for attention are present in the Attention Index score but not as prominent as those in the Screening Attention Domain score. Similarly, the practice effects for memory are present in the Memory Index score but are not as dramatic as those in the Screening Memory Domain score. Additionally, the memory practice effects as a function of age are not as prominent for the Memory Index. This trend is likely due to the increased amount of information the examinee is required to learn in the Memory Module tests as compared to the requirements of the Screening Memory Domain tests. Regardless, for individuals who are retested, it is always prudent for the examiner to consider practice effects in the interpretation of test scores.

Table 5.17 presents the descriptive statistics and stability coefficients of the Screening Domain, Total Screening Index, module index, and Total NAB Index scores by NAB Form. Similar to the previously discussed practice effects findings, the Screening Memory Domain score and Memory Index show comparable practice effects. In this study, Form 2 participants tended to have larger practice effects on the Screening Spatial Domain score as compared to Form 1 participants. Overall, for the Total Screening Index and the Total NAB Index, there were negligible differences in practice effects between test forms. It is also important to note that differences in practice effects between NAB test forms were generally within the expected standard error of measurement.

As previously noted, many of the NAB primary, secondary, and descriptive scores have limited variability in normal populations, which has the effect of attenuating traditional stability measures (i.e., reducing the Pearson correlation). Attenuation has the effect of significantly underestimating the “true” temporal stability of the measure. Although there are statistical corrections for attenuation, they are not without controversy and limitations (see Nunnally & Bernstein, 1994). As previously discussed, a number of the stability coefficients presented in Tables 5.7 through 5.12 are quite low, and the user would be advised to consider a number of conceptual (i.e., expected trait normality in nonimpaired

individuals) and psychometric (e.g., score attenuation) considerations when interpreting the stability of these measures. In general, the stability coefficients presented for the NAB primary, secondary, and descriptive scores are commensurate with similar neuropsychological measures in current use that have similar applications and clinical utility.

Equivalent Forms Reliability

The reliability of the NAB was also evaluated according to generalizability theory (Brennan, 2001; Cronbach, Gleser, Nanda, & Rajaratnam, 1972; Shavelson & Webb, 1991). In contrast to classical psychometric theory that posits true scores and a unitary, or undifferentiated, source of error, the application of generalizability theory allows the partitioning of various sources of variance with the familiar analysis of variance (ANOVA) design. The purpose of this study was to evaluate the reliability of NAB scores as a function of test form. Generalizability coefficients are considered analogues of traditional reliability estimates. However, in contrast to the magnitude of traditional reliability estimates, generalizability coefficients of .60 or higher are regarded as demonstrating very good reliability (Cicchetti & Sparrow, 1981; Mitchell, 1979).

The generalizability study sample consisted of 100 participants who were administered NAB Forms 1 and 2 in a counterbalanced design. The average Form 1–Form 2 test–retest interval was 25.0 days (*SD* = 6.3 days). The participants ranged in age from 18 to 84 years, with a mean age of 57.9 years (*SD* = 19.5 years). The percentages of the sample by education level were 26.0% ≤11 years, 24.0% 12 years, 30.0% 13 to 15 years, and 20.0% ≥16 years, with a mean level of education of 13.0 years (*SD* = 2.5 years). The sample consisted of 66.0% female and 34.0% male participants and the following proportions of race/ethnicity: 82.0% Caucasian, 6.0% African American, 2.0% Hispanic, and 10.0%, Other race/ethnicity.

The generalizability (*G*) coefficients for the NAB scores are shown in Tables 5.18 through 5.23. As shown in these tables, the *G* coefficients for the primary scores are listed along with the standard error of measurement for each. Additionally, the *G* coefficient is listed by module for a number of other appropriate supplementary and descriptive scores. The median *G* coefficient for the Screening Module primary scores is .75. The following median *G* coefficients for the primary scores were obtained: .80 for the Attention Module, .62 for the Language Module, .72 for the Memory Module, .73 for the Spatial Module, and .83 for the Executive Functions Module. The median *G* coefficients for the module primary scores all exceed .60 and, overall, demonstrate good to very good reliability.

Table 5.18
Summary Reliability Coefficients and Standard Errors of Measurement (SEMs)
for Screening Module Primary, Secondary, and Descriptive Scores for Forms 1 and 2

Test	Acronym	Alpha coefficient^a	SEM based on alpha coefficient	Stability coefficient^b	SEM based on stability coefficient	G^d	SEM based on G
Primary score							
Screening Digits Forward	S-DGF	.78	4.69	.65	5.91	.87	3.56
Screening Digits Backward	S-DGB	.79	4.58	.67	5.78	.88	3.51
Screening Numbers & Letters Part A Speed	S-N&L _A -spd			.69	5.53	.88	3.44
Screening Numbers & Letters Part A Errors	S-N&L _A -err			.35	8.07	.41	7.70
Screening Numbers & Letters Part A Efficiency	S-N&L _A -eff			.71	5.38	.90	3.19
Screening Numbers & Letters Part B Efficiency	S-N&L _B -eff			.53	6.85	.75	4.99
Screening Auditory Comprehension	S-AUD	.48		.43	7.58	.45	7.39
Screening Naming	S-NAM	.36	8.01	.33	8.16	.46	7.33
Screening Shape Learning Immediate Recognition	S-SHL-irg		5.29	.11	9.42	.53	6.89
Screening Shape Learning Delayed Recognition	S-SHL-drg			.36	7.98	.33	8.17
Screening Story Learning Immediate Recall	S-STL-irc	.69	5.56	.42	7.61	.74	5.07
Screening Story Learning Delayed Recall	S-STL-drc	.72	5.32	.43	7.57	.81	4.39
Screening Visual Discrimination	S-VIS	.24	8.00	.16	9.16	.59	6.38
Screening Design Construction	S-DES	.31	8.72	.58	6.49	.73	5.17
Screening Mazes	S-MAZ	.55	6.69	.63	6.07	.91	2.95
Screening Word Generation	S-WGN			.63	6.10	.76	4.93
Secondary score							
Screening Orientation	S-ORN			% agreement ^c			
Screening Digits Forward Longest Span	S-DGF-spn			89.0%		.79	
				95.6%			

(continued)

Table 5.18 (continued)
Summary Reliability Coefficients and Standard Errors of Measurement (SEMs)
for Screening Module Primary, Secondary, and Descriptive Scores for Forms 1 and 2

Test	Acronym	Alpha coefficient^a	SEM based on alpha coefficient	Stability coefficient^b	SEM based on stability coefficient	G^d	SEM based on G
Screening Digits Backward Longest Span	S-DGB-spn			90.3%		.80	
Screening Story Learning Percent Retention	S-STL-%rt			82.7%		.47	
Screening Shape Learning Percent Retention	S-SHL-%rt			83.7%			
Screening Word Generation Perseverations	S-WGN-psv			75.7%			
Descriptive score							
Screening Orientation to Self	S-ORN-slf			100.0%			
Screening Orientation to Time	S-ORN-tim			94.3%			
Screening Orientation to Place	S-ORN-plc			100.0%			
Screening Orientation to Situation	S-ORN-sit			100.0%			
Screening Auditory Comprehension Colors	S-AUD-col			100.0%			
Screening Auditory Comprehension Shapes	S-AUD-shp			98.3%			
Screening Auditory Comprehension Colors/Shapes/Numbers	S-AUD-csn			99.2%		.44	
Screening Naming Percent Correct After Semantic Cuing	S-NAM-sem%			65.5%		.34	
Screening Naming Percent Correct After Phonemic Cuing	S-NAM-pho%			79.6%		.59	

Note. *Ns* = 30 for interrater reliability study and 100 for generalizability study (ages 18-84 years). The relevant reliability coefficients and the population standard deviation (i.e., 10) were used to compute the standard errors of measurement. Average interrater coefficient percent agreements are expressed as percentages, all others as ICCs.

^aArithmetic average of reliability of Forms 1 and 2 across all ages. ^bArithmetic average of stability coefficient of Forms 1 and 2 across all ages. ^cReliability of secondary and descriptive scores is reported as average percentage agreement. ^dGeneralizability coefficient.

Table 5.19
Summary of Reliability Coefficients and Standard Errors of Measurement (SEMs)
for Attention Module Primary, Secondary, and Descriptive Scores for Forms 1 and 2

Test	Acronym	Alpha coefficient^a	SEM based on alpha coefficient	Stability coefficient^b	SEM based on stability coefficient	G^d	SEM based on G
Primary score							
Digits Forward	DGF	.78	4.69	.65	5.91	.87	3.56
Digits Backward	DGB	.79	4.58	.67	5.78	.88	3.51
Dots	DOT			.44	7.51	.80	4.49
Numbers & Letters Part A Speed	N&L _A -spd			.86	3.70	.91	3.07
Numbers & Letters Part A Errors	N&L _A -err			.56	6.60	.75	5.05
Numbers & Letters Part A Efficiency	N&L _A -eff			.87	3.65	.92	2.88
Numbers & Letters Part B Efficiency	N&L _B -eff			.58	6.50	.73	5.19
Numbers & Letters Part C Efficiency	N&L _C -eff			.58	6.48	.66	5.84
Numbers & Letters Part D Efficiency	N&L _D -eff			.65	5.94	.69	5.54
Numbers & Letters Part D Disruption	N&L _D -dis			.47	7.25	.34	8.12
Driving Scenes	DRV			.64	6.00	.89	3.32
Secondary score							
Orientation	ORN			% agreement ^c			
Digits Forward Longest Span	DGF-spn			89.0%		.79	4.54
Digits Backward Longest Span	DGB-spn			95.6%		.80	4.53
Descriptive score							
Orientation to Self	ORN-slf			100.0%			
Orientation to Time	ORN-tim			94.3%			
Orientation to Place	ORN-plc			100.0%			
Orientation to Situation	ORN-sit			100.0%			

Note. *Ns* = 30 for interrater reliability study and 100 for generalizability study (ages 18-84 years). The relevant reliability coefficients and the population standard deviation (i.e., 10) were used to compute the standard errors of measurement. Average interrater coefficient percent agreements are expressed as percentages, all others as ICCs.

^aArithmetic average of reliability of Forms 1 and 2 across all ages. ^bArithmetic average of stability coefficient of Forms 1 and 2 across all ages. ^cReliability of secondary and descriptive scores is reported as average percentage agreement. ^dGeneralizability coefficient.

Table 5.20
Summary of Reliability Coefficients and Standard Errors of Measurement (SEMs)
for Language Module Primary, Secondary, and Descriptive Scores for Forms 1 and 2

Test	Acronym	Alpha coefficient^a	SEM based on alpha coefficient	Stability coefficient^b	SEM based on stability coefficient	G^d	SEM based on G
Primary score							
Oral Production	OPD	.84	4.00	.23	8.78	.71	5.42
Auditory Comprehension	AUD	.48	7.21	.26	8.58	.54	6.77
Naming	NAM	.76	4.90	.70	5.43	.72	5.25
Writing	WRT			.24	8.74	.60	6.32
Bill Payment	BIL	.69	5.57	.41	7.65	.62	6.16
Secondary score							
Reading Comprehension	RCN			% agreement ^c			
Writing Legibility	WRT-leg			99.2%			
Writing Spelling	WRT-spl			100.0%		.80	4.47
Writing Syntax	WRT-syn			98.3%			
Writing Conveyance	WRT-cnv			100.0%		.74	5.12
				88.6%			
Descriptive score							
Auditory Comprehension Colors	AUD-col			100.0%			
Auditory Comprehension Shapes	AUD-shp			98.3%			
Auditory Comprehension Colors/Shapes/Numbers	AUD-csn			99.2%		.44	
Auditory Comprehension Pointing	AUD-pnt			100.0%			
Auditory Comprehension Yes/No	AUD-y/n			96.1%		.24	
Auditory Comprehension Paper Folding	AUD-flt			90.7%		.18	
Naming Percent Correct After Semantic Cuing	NAM-sem%			79.4%		.44	
Naming Percent Correct After Phonemic Cuing	NAM-pho%			75.3%		.47	
Reading Comprehension Words	RCN-wrd			99.2%			
Reading Comprehension Sentences	RCN-sen			97.8%			

Note. *Ns* = 30 for interrater reliability study and 100 for generalizability study (ages 18-84 years). The relevant reliability coefficients and the population standard deviation (i.e., 10) were used to compute the standard errors of measurement. Average interrater coefficient percent agreements are expressed as percentages, all others as ICCs.

^aArithmetic average of reliability of Forms 1 and 2 across all ages. ^bArithmetic average of stability coefficient of Forms 1 and 2 across all ages. ^cReliability of secondary and descriptive scores is reported as average percentage agreement. ^dGeneralizability coefficient.

Table 5.21
Summary of Reliability Coefficients and Standard Errors of Measurement (SEMs)
for Memory Module Primary, Secondary, and Descriptive Scores for Forms 1 and 2

Test	Acronym	Alpha coefficient^a	SEM based on alpha coefficient	Stability coefficient^b	SEM based on stability coefficient	G^d	SEM based on G
Primary score							
List Learning List A Immediate Recall	LLA-irc			.47	7.26	.80	4.47
List Learning List B Immediate Recall	LLB-irc			.48	7.24	.56	6.62
List Learning List A Short Delayed Recall	LLA-sd:drc			.44	7.46	.71	5.38
List Learning List A Long Delayed Recall	LLA-ls:drc			.58	6.50	.67	5.73
Shape Learning Immediate Recognition	SHL-irg	.69	5.60	.43	7.58	.73	5.23
Shape Learning Delayed Recognition	SHL-drg	.47	7.30	.41	7.70	.74	5.09
Story Learning Phrase Unit Immediate Recall	STL-irc:phu	.86	3.69	.52	6.91	.53	6.83
Story Learning Phrase Unit Delayed Recall	STL-drc:phu	.86	3.77	.61	6.25	.45	7.43
Daily Living Memory Immediate Recall	DLM-irc			.52	6.96	.85	3.94
Daily Living Memory Delayed Recall	DLM-drc			.47	7.27	.75	5.04
Secondary score							
				% agreement^c			
List Learning List A Trial 1 Immediate Recall	LLA1-irc			91.6%		.72	
List Learning List A Trial 2 Immediate Recall	LLA2-irc			89.8%		.80	
List Learning List A Trial 3 Immediate Recall	LLA3-irc			93.4%		.68	
List Learning List A Percent Retention	LLA-%rt			85.8%		.24	

(continued)

Table 5.21 (continued)
Summary of Reliability Coefficients and Standard Errors of Measurement (SEMs)
for Memory Module Primary, Secondary, and Descriptive Scores for Forms 1 and 2

Test	Acronym	Alpha coefficient^a	SEM based on alpha coefficient	Stability coefficient^b	SEM based on stability coefficient	G^d	SEM based on G
List Learning List A Long Delayed Forced-Choice Recognition	LLA-dfc			91.6%			
List Learning List A Long Delayed Forced-Choice Recognition False Alarms	LLA-fa			88.5%		.59	
List Learning List A Discriminability Index	LLA-dis			94.4%		.55	
List Learning List A Recall vs. Recognition	LLA-rvr			90.3%		.61	
Shape Learning Trial 1 Immediate Recognition	SHL1-irg			93.3%		.52	
Shape Learning Trial 2 Immediate Recognition	SHL2-irg			85.8%		.75	
Shape Learning Trial 3 Immediate Recognition	SHL3-irg			92.9%		.53	
Shape Learning Percent Retention	SHL-%rt			72.4%		.13	
Shape Learning Delayed Forced-Choice Recognition	SHL-dfc			89.0%		.39	
Shape Learning Delayed Forced-Choice Recognition False Alarms	SHL-fa			88.9%		.50	
Shape Learning Discriminability Index	SHL-dis			95.1%		.36	
Story Learning Trial 1 Phrase Unit	STL1-irc:phu			83.8%		.49	
Story Learning Trial 2 Phrase Unit	STL2-irc:phu			90.7%		.46	
Story Learning Thematic Unit Immediate Recall	STL-irc:thu			92.9%		.54	
Story Learning Trial 1 Thematic Unit	STL1-irc:thu			91.3%		.44	
Story Learning Trial 2 Thematic Unit	STL2-irc:thu			98.7%		.49	
Story Learning Thematic Unit Delayed Recall	STL-drc:thu			96.5%		.50	

(continued)

Table 5.21 (continued)
Summary of Reliability Coefficients and Standard Errors of Measurement (SEMs)
for Memory Module Primary, Secondary, and Descriptive Scores for Forms 1 and 2

Test	Acronym	Alpha coefficient^a	SEM based on alpha coefficient	Stability coefficient^b	SEM based on stability coefficient	G^d	SEM based on G
Story Learning Phrase Unit Percent Retention	STL-%rt			88.5%		.18	
Daily Living Memory Retention	DLM-rt			89.5%		.44	
Daily Living Memory Delayed Recognition	DLM-drg			87.7%		.40	
Daily Living Memory Recall vs. Recognition	DLM-rvr			91.8%		.13	
Medication Instructions Immediate Recall	MED-irc			92.5%		.81	
Medication Instructions Delayed Recall	MED-drc			93.0%		.35	
Medication Instructions Delayed Recognition	MED-drg			85.9%		.02	
Name/Address/Phone Immediate Recall	NAP-irc			93.0%		.75	
Name/Address/Phone Delayed Recall	NAP-drc			95.7%		.73	
Name/Address/Phone Delayed Recognition	NAP-drg			93.4%		.44	
Descriptive score							
List Learning Semantic Clusters	LL-sem			90.8%		.72	
List Learning Perseverations	LL-psv			68.4%		.22	
List Learning Intrusions	LL-int			83.3%		.11	

Note. $N_s = 30$ for interrater reliability study and 100 for generalizability study (ages 18-84 years). The relevant reliability coefficients and the population standard deviation (i.e., 10) were used to compute the standard errors of measurement. Average interrater coefficient percent agreements are expressed as percentages, all others as ICCs.

^aArithmetic average of reliability of Forms 1 and 2 across all ages. ^bArithmetic average of stability coefficient of Forms 1 and 2 across all ages. ^cReliability of secondary and descriptive scores is reported as average percentage agreement. ^dGeneralizability coefficient.

Table 5.22
Summary of Reliability Coefficients and Standard Errors of Measurement (SEMs)
for Spatial Module Primary and Secondary Scores for Forms 1 and 2

Test	Acronym	Alpha coefficient^a	SEM based on alpha coefficient	Stability coefficient^b	SEM based on stability coefficient	G^d	SEM based on G
Primary score							
Visual Discrimination	VIS	.67	5.74	.38	7.90	.68	5.69
Design Construction	DES	.67	5.74	.68	5.63	.92	2.88
Figure Drawing Copy	FGD –cpy			.13	9.31	.77	4.80
Figure Drawing Copy Organization	FGD –cpy:org			.35	8.09	.38	7.89
Figure Drawing Immediate Recall	FGD –irc			.54	6.82	.73	5.17
Map Reading	MAP	.65	5.92	.54	6.79	.73	5.22
Secondary score							
				% agreement^c			
Figure Drawing Copy Fragmentation	FGD –cpy:frg			0%		.25	
Figure Drawing Copy Planning	FGD –cpy:pln			0%		.27	
Figure Drawing Immediate Recall Organization	FGD –irc:org			0%		.61	
Figure Drawing Immediate Recall Fragmentation	FGD –irc:frg			0%		.56	
Figure Drawing Immediate Recall Planning	FGD –irc:pln			0%		.44	
Figure Drawing Percent Retention	FGD –%rt			0%		.63	

Note. *Ns* = 30 for interrater reliability study and 100 for generalizability study (ages 18-84 years). The relevant reliability coefficients and the population standard deviation (i.e., 10) were used to compute the standard errors of measurement. Average interrater coefficient percent agreements are expressed as percentages, all others as ICCs.

^aArithmetic average of reliability of Forms 1 and 2 across all ages. ^bArithmetic average of stability coefficient of Forms 1 and 2 across all ages. ^cReliability of secondary scores are reported as average percentage agreement. ^dGeneralizability coefficient.

Table 5.23
Summary of Reliability Coefficients and Standard Errors of Measurement (SEMs)
for Executive Functions Module Primary and Secondary Scores for Forms 1 and 2

Test	Acronym	Alpha coefficient^a	SEM based on alpha coefficient	Stability coefficient^b	SEM based on stability coefficient	G^d	SEM based on G
Primary score							
Mazes	MAZ	.77	4.80	.57	6.55	.95	2.35
Judgment	JDG	.45	7.42	.43	7.57	.60	6.30
Categories	CAT			.54	6.81	.89	3.36
Word Generation	WGN			.64	5.97	.77	4.80
Secondary score							
Word Generation Perseverations	WGN-psv			% agreement^c			
				82.4%		.51	

Note. *Ns* = 30 for interrater reliability study and 100 for generalizability study (ages 18-84 years). The relevant reliability coefficients and the population standard deviation (i.e., 10) were used to compute the standard errors of measurement. Average interrater coefficient percent agreements are expressed as percentages, all others as ICCs.

^aArithmetic average of reliability of Forms 1 and 2 across all ages. ^bArithmetic average of stability coefficient of Forms 1 and 2 across all ages. ^cReliability of secondary scores is reported as average percentage agreement. ^dGeneralizability coefficient.

It is recognized that NAB users may need to evaluate various aspects of score reliability depending upon the referral question and questions of interest. Therefore, Tables 5.18 through 5.23 summarize the various estimates of reliability including alpha coefficients, test–retest reliability, temporal stability, percentage agreement, interrater reliability, and *G* coefficients. Where appropriate, the standard error of measurement corresponding to the relevant reliability coefficient is calculated and included to the right of the reliability coefficient.

Interrater Agreement

The consistency of agreement of test scores from rater to rater is also an important indication of a test’s reliability, especially for those subtests that require scorer judgment and decision making. Thirty Form 1 and 30 Form 2 standardization protocols were randomly selected and independently scored by two raters. The interrater reliability for the NAB was examined for the following subtests: Writing, Story Learning, Figure Drawing, Judgment, and Categories (see Table 5.24).

Two experienced standardization examiners served as raters for the Writing subtest. Due to reduced variability

between raters, average percentage agreement coefficients were calculated between the raters. The average interrater agreement percentages for NAB Forms 1 and 2 Writing scores are shown in Table 5.24. The percentage agreement ranged from 95.0% for the primary Writing score (WRT) to 100.0% for Writing Spelling (WRT–spl). Overall, there was an average 98.1% interrater agreement for the Writing scores.

For the remaining subtests (i.e., Story Learning, Figure Drawing, Judgment, and Categories), one-way single-measure intraclass correlation coefficients (ICC) were calculated (see Table 5.24). The average ICCs for Story Learning range from .93 for Story Learning Thematic Unit Immediate Recall (STL–irc:thu) to .99 for Story Learning Phrase Unit Immediate Recall (STL–irc:phu). The average ICCs for Figure Drawing range from .83 for Figure Drawing Copy (FGD–cpy) to .96 for Figure Drawing Copy Fragmentation (FGD–cpy:frg). The average ICCs for Judgment and Categories were .85 and .97, respectively. The interrater reliability study demonstrated overall excellent reliability, and these data show that, for the NAB tests that require scoring judgment, scoring can be accomplished reliably by appropriately trained scorers.

Table 5.24
Summary of Interrater Reliability
for Selected Scores for Forms 1 and 2

Test	Acronym	Average interrater reliability (ICC)/ % agreement ^a
Writing	WRT	95.0%
Writing Syntax	WRT–syn	95.5%
Writing Legibility	WRT–leg	98.5%
Writing Spelling	WRT–spl	100.0%
Writing Conveyance	WRT–cnv	98.5%
Story Learning Trial 1 Phrase Unit	STL–irc:phu	.99
Story Learning Trial 1 Thematic Unit	STL–irc:thu	.93
Figure Drawing Copy	FGD–cpy	.83
Figure Drawing Copy Organization	FGD–cpy:org	.95
Figure Drawing Copy Fragmentation	FGD–cpy:frg	.96
Figure Drawing Copy Planning	FGD–cpy:pln	.88
Figure Drawing Immediate Recall	FGD–irc	.90
Figure Drawing Immediate Recall Organization	FGD–irc:org	.93
Figure Drawing Immediate Recall Fragmentation	FGD–irc:frg	.89
Figure Drawing Immediate Recall Planning	FGD–irc:pln	.87
Judgment	JDG	.85
Categories	CAT	.97

Note. *N* = 30.

^a% agreements are reported as percentages; interrater reliability is reported as an intraclass correlation coefficient (ICC).

Reliability of the NAB Composite Scores

Given that alpha coefficient is an inappropriate estimate for many NAB scores, *G* coefficients were uniformly used as reliability estimates for the purpose of calculating the reliability estimates of Screening Domain, Total Screening Index, module index, and Total NAB Index scores. The reliability coefficients for all composite scores were calculated with the formula recommended by Guilford (1954) and Nunnally (1978). Table 5.25 presents the composite reliabilities for the Screening Domain, Total Screening Index, module index, and Total NAB Index scores, along with the corresponding standard errors of measurement (*SEMs*). The reliabilities for the Screening Domain scores range from .55 for the Screening Language Domain score to .91 for the Screening Attention Domain score. The median Screening Domain score reliability coefficient is .79, and the Total Screening Index reliability coefficient is .80. The module index score reliabilities range from .79 for the Language Index score to .93 for both the Attention Index and Memory Index scores. The reliability coefficient of the Total NAB Index score is .96. As expected, the module index scores generally have higher reliabilities than the corresponding Screening Domain scores. These data clearly demonstrate that the obtained reliability estimates range from good to excellent. Additionally, given the psychometric issues

discussed previously (e.g., restriction of score range) and the unique characteristics of many of the neuropsychological constructs measured, the overall NAB reliability estimates are excellent.

STANDARD ERRORS OF MEASUREMENT AND CONFIDENCE INTERVALS

The standard error of measurement (*SEM*) is an estimate of the expected error in an individual's test score. As can be seen in the following formula, the standard error of measurement is inversely related to the reliability of a score:

$$SEM = SD * \text{SQRT}(1-r_{xx})$$

In this formula, *SD* is the standard deviation of the obtained *T* score (i.e., 10) or the standard deviation of the obtained Screening Domain, Total Screening Index, module index, and Total NAB Index scores (i.e., 15). For the various reliability estimates presented in Tables 5.18 through 5.23, the corresponding *SEM* was computed with this formula. As is obvious from the formula, the smaller the *SEM* (i.e., measurement error), the more confidence the user can have in the precision of the obtained test score. When a confidence interval is developed for a particular score, the *SEM* is used to center the interval around the obtained score.

Table 5.25
Summary of Reliability Coefficients and Standard Errors of Measurement (*SEMs*) for Screening Domain, Total Screening Index, Module Index, and Total NAB Index Scores

Domain/Index score	Reliability (<i>G</i>)	<i>SEM</i>
Screening Module Domain/Index		
Screening Attention Domain	.91	4.50
Screening Language Domain	.55	10.06
Screening Memory Domain	.79	6.87
Screening Spatial Domain	.71	8.08
Screening Executive Functions Domain	.86	5.61
Total Screening Index	.80	6.71
Index score		
Attention Index	.93	3.97
Language Index	.79	6.87
Memory Index	.93	3.97
Spatial Index	.88	5.20
Executive Functions Index	.89	4.97
Total NAB Index	.96	3.00

Note. The reliabilities for the indexes were calculated with the formula recommended by Guilford (1954) and Nunnally (1978). The *SEMs* are reported in *T*-score units. The relevant reliability coefficients and the population standard deviation (i.e., 15) were used to compute the *SEMs*.

For example, if an individual obtains a particular T score of 50 and the SEM is 5.10, using a 95% certainty level (i.e., $\pm 1.96 SEMs$) the user can be 95% certain that the individual's true score is in the T -score range of 40 to 60.

When evaluating any particular primary, secondary, or descriptive NAB score, the user must be mindful of the various factors that can influence reliability coefficients and, in turn, influence the $SEMs$ and resulting computed confidence intervals. For example, reliability estimates that are attenuated due to restriction of score range will have relatively larger confidence intervals. In this example, these relatively larger confidence intervals may very well be an artifact of the inherent psychometric limitations that produced the reliability coefficient and, thus, contributed to the computation of the larger confidence interval.

For most users, however, the Screening Domain, Total Screening Index, module index, and Total NAB Index scores will be the primary focus of interpretation and, therefore, confidence intervals have been developed for the user at the 90% and 95% levels. These are presented in the normative tables of the selected norms manual.

SCORE DIFFERENCES

An important consideration in the interpretation of an individual examinee's performance is the magnitude of difference between scores. Score differences or discrepancies have at least two important aspects. The first issue specifically addresses the concept of measurement error. That is to say, after consideration of measurement error, are statistically reliable score differences present? The second issue addresses the actual occurrence rates as expressed in cumulative percentages in such discrepancies that are actually present in the population. It is quite possible to obtain difference scores that are statistically significant but that occur relatively frequently in the standardization sample and, by extrapolation, in the overall population.

Statistical Significance of Module Primary Scores and Index Score Differences

The probability level associated with score differences is a reflection of the reliability of such differences. When the probability is very low (e.g., $p < .01$), the test user can have a high degree of confidence that this difference is reliable. The standard error of measurement of score differences (SEM_{diff}) formula is used as a basis in making this determination. Multiplying the SEM_{diff} by the desired z value provides the critical value that the score differences must achieve to be considered statistically significant.

The differences required for statistical significance at the $p = .01$ and $p = .05$ levels between all combinations of Screening Domain scores and the Total Screening Index score are presented in Table 5.26. Table 5.27 presents the analogous values required for statistical significance for all combinations of the module indexes and Total NAB Index. In Table 5.27, for example, the difference between the Attention Index and Language Index must be at least 15.6 (i.e., 16) points in order to be statistically different at the $p = .05$ level, and at least 20.5 (i.e., 21) points at the $p = .01$ level. Additionally, Tables 5.28 through 5.33 present the required statistical differences between the primary scores for each of the six NAB modules. Critical values for statistically significant differences are given for both the $p = .05$ and $p = .01$ levels. In Table 5.28, for example, a difference between Screening Digits Forward (S-DGF) and Screening Digits Backward (S-DGB) of at least 9.8 (i.e., 10) T -score points is required at the $p = .05$ level (12.9 points at the $p = .01$ level) for statistical difference. Because all of the generalizability study participants in the demographically corrected standardization sample are also contained in the U.S. Census-matched standardization sample, Tables 5.28 through 5.33 provide the required critical values for evaluating primary score statistical differences between primary T scores based on either standardization sample.

Base Rates of Index Score Differences

The base rate refers to the cumulative percentage of the discrepancy occurrence that is found in the standardization sample, and, by extrapolation, in the general population. As previously noted, statistically significant differences between an individual's Screening domain or module index scores may or may not be rare in the general population. Therefore, it is important to evaluate statistically significant differences within the context of population rarity: A statistically significant score difference within the context of a low base rate may be clinically important, while such a difference with a relatively higher base rate would not necessarily warrant the same clinical significance.

The frequencies of score differences between Screening Domain and Total Screening Index scores are presented in Tables 5.34 and 5.35 for the demographically corrected standardization sample and the age-based, U.S. Census-matched sample, respectively. The frequencies of score differences between main module index scores and the Total NAB Index score are presented in Tables 5.36 and 5.37 for the demographically corrected standardization sample and the age-based, U.S. Census-matched sample, respectively. These tables present the cumulative percentages of the absolute differences (i.e., regardless of the direction of the score difference) between each possible pair of scores.

Table 5.26
Discrepancies Between Screening Module Domain/Index Scores Required
for Statistical Significance at the .05 and .01 Levels of Confidence

Module Domain/Index pair	Acronym	Required discrepancy	
		$p = .05$	$p = .01$
Screening Attention Domain – Screening Language Domain	S-ATT – S-LAN	21.6	28.4
Screening Attention Domain – Screening Memory Domain	S-ATT – S-MEM	16.1	21.2
Screening Attention Domain – Screening Spatial Domain	S-ATT – S-SPT	18.1	23.9
Screening Attention Domain – Screening Executive Functions Domain	S-ATT – S-EXE	14.1	18.6
Screening Attention Domain – Total Screening Index	S-ATT – S-NAB	15.8	20.8
Screening Language Domain – Screening Memory Domain	S-LAN – S-MEM	23.9	31.4
Screening Language Domain – Screening Spatial Domain	S-LAN – S-SPT	25.3	33.3
Screening Language Domain – Screening Executive Functions Domain	S-LAN – S-EXE	22.6	29.7
Screening Language Domain – Total Screening Index	S-LAN – S-NAB	23.7	31.2
Screening Memory Domain – Screening Spatial Domain	S-MEM – S-SPT	20.8	27.4
Screening Memory Domain – Screening Executive Functions Domain	S-MEM – S-EXE	17.4	22.9
Screening Memory Domain – Total Screening Index	S-MEM – S-NAB	18.8	24.8
Screening Spatial Domain – Screening Executive Functions Domain	S-SPT – S-EXE	19.3	25.4
Screening Spatial Domain – Total Screening Index	S-SPT – S-NAB	20.6	27.1
Screening Executive Functions Domain – Total Screening Index	S-EXE – S-NAB	17.1	22.6

Note. Differences required for statistical significance are based on the standard errors of measurement (*SEMs*) of each index and are calculated with the following formula: Difference score = $Z\sqrt{(SEM_a^2 + SEM_b^2)}$, where the *Z* is the normal curve value associated with the .05 and .01 significance levels and SEM_a and SEM_b are the standard errors of measurement of the two domains/indexes.

Table 5.27
Discrepancies Between Module Index and Total NAB Index Scores Required
for Statistical Significance at the .05 and .01 Levels of Confidence

Module Index pair	Acronym	Required discrepancy	
		$p = .05$	$p = .01$
Attention Index – Language Index	ATT – LAN	15.6	20.5
Attention Index – Memory Index	ATT – MEM	11.0	14.5
Attention Index – Spatial Index	ATT – SPT	12.8	16.9
Attention Index – Executive Functions Index	ATT – EXE	12.5	16.4
Attention Index – Total NAB Index	ATT – T-NAB	9.8	12.8
Language Index – Memory Index	LAN – MEM	15.6	20.5
Language Index – Spatial Index	LAN – SPT	16.9	22.2
Language Index – Executive Functions Index	LAN – EXE	16.6	21.9
Language Index – Total NAB Index	LAN – T-NAB	14.7	19.4
Memory Index – Spatial Index	MEM – SPT	12.8	16.9
Memory Index – Executive Functions Index	MEM – EXE	12.5	16.4
Memory Index – Total NAB Index	MEM – T-NAB	9.8	12.8
Spatial Index – Executive Functions Index	SPT – EXE	14.1	18.6
Spatial Index – Total NAB Index	SPT – T-NAB	11.8	15.5
Executive Functions Index – Total NAB Index	EXE – T-NAB	11.4	15.0

Note. Differences required for statistical significance are based on the standard errors of measurement (*SEMs*) of each index and are calculated with the following formula: Difference score = $Z\sqrt{(SEM_a)^2 + (SEM_b)^2}$, where the *Z* is the normal curve value associated with the .05 and .01 significance levels and SEM_a and SEM_b are the standard errors of measurement of the two indexes.

Table 5.28
Discrepancies Between Screening Module Primary Scores Required for
Statistical Significance at the .05 and .01 Levels of Confidence for All Ages

Primary score	Acronym	S-DGF	S-DGB	S-N&L _A -spd	S-N&L _A -err	S-N&L _A -eff	S-N&L _B -eff	S-AUD	S-NAM	S-SHL-irg	S-SHL-drg	S-STL-irg	S-STL-drc	S-VIS	S-DES	S-MAZ	S-WGN
Screening Digits Forward	S-DGF	—	9.8	9.7	16.6	9.4	12.0	16.1	16.0	15.2	17.5	12.1	11.1	14.3	12.3	9.1	11.9
Screening Digits Backward	S-DGB	12.9	—	9.6	16.6	9.3	12.0	16.0	15.9	15.2	17.4	12.1	11.0	14.3	12.2	9.0	11.9
Screening Numbers & Letters Part A Speed	S-N&L _A - spd	12.8	12.7	—	16.5	9.2	11.9	16.0	15.9	15.1	17.4	12.0	10.9	14.2	12.2	8.9	11.8
Screening Numbers & Letters Part A Errors	S-N&L _A - err	21.9	21.8	21.8	—	16.3	18.0	20.9	20.8	20.3	22.0	18.1	17.4	19.6	18.2	16.2	17.9
Screening Numbers & Letters Part A Efficiency	S-N&L _A - eff	12.3	12.2	12.1	21.5	—	11.6	15.8	15.7	14.9	17.2	11.7	10.6	14.0	11.9	8.5	11.5
Screening Numbers & Letters Part B Efficiency	S-N&L _B - eff	15.8	15.7	15.6	23.7	15.3	—	17.5	17.4	16.7	18.8	13.9	13.0	15.9	14.1	11.4	13.7
Screening Auditory Comprehension	S-AUD	21.2	21.1	21.0	27.5	20.8	23.0	—	20.4	19.8	21.6	17.6	16.8	19.1	17.7	15.6	17.4
Screening Naming	S-NAM	21.0	21.0	20.9	27.4	20.6	22.9	26.8	—	19.7	21.5	17.5	16.7	19.0	17.6	15.5	17.3
Screening Shape Learning Immediate Recognition	S-SHL- irg	20.0	20.0	19.9	26.7	19.6	22.0	26.1	26.0	—	20.9	16.8	16.0	18.4	16.9	14.7	16.6
Screening Shape Learning Delayed Recognition	S-SHL- drg	23.0	22.9	22.9	29.0	22.6	24.7	28.4	28.3	27.6	—	18.8	18.2	20.3	18.9	17.0	18.7
Screening Story Learning Immediate Recall	S-STL- irc	16.0	15.9	15.8	23.8	15.5	18.4	23.1	23.0	22.1	24.8	—	13.1	16.0	14.2	11.5	13.9
Screening Story Learning Delayed Recall	S-STL- drc	14.6	14.5	14.4	22.9	14.0	17.2	22.2	22.0	21.1	23.9	17.3	—	15.2	13.3	10.4	12.9
Screening Visual Discrimination	S-VIS	18.9	18.8	18.7	25.8	18.4	20.9	25.2	25.1	24.2	26.7	21.0	20.0	—	16.1	13.8	15.8
Screening Design Construction	S-DES	16.2	16.1	16.0	23.9	15.7	18.5	23.3	23.1	22.2	24.9	18.7	17.5	21.2	—	11.7	14.0
Screening Mazes	S-MAZ	11.9	11.8	11.7	21.3	11.2	15.0	20.5	20.4	19.3	22.4	15.1	13.7	18.1	15.4	—	11.3
Screening Word Generation	S-WGN	15.7	15.6	15.5	23.6	15.2	18.1	22.9	22.8	21.9	24.6	18.2	17.0	20.8	18.4	14.8	—

Note. Differences required for statistical significance at the .05 level appear above the diagonal in the shaded area; differences significant at the .01 level appear below the diagonal. The differences required for statistical significance were calculated with the following formula: Difference score = $Z\sqrt{(SEM_a^2 + SEM_b^2)}$ where the Z is the normal curve value associated with the desired significance level and SEM_a and SEM_b are the standard errors of measurement of the scores.

Table 5.29
Discrepancies Between Attention Module Primary Scores Required for
Statistical Significance at the .05 and .01 Levels of Confidence for All Ages

Primary score	Acronym	DGF	DGB	DOT	N&L _A -spd	N&L _A -err	N&L _A -eff	N&L _B -eff	N&L _C -eff	N&L _D -eff	N&L _D -dis	DRV
Digits Forward	DGF	—	9.8	11.2	9.2	12.1	9.0	12.3	13.4	12.9	17.4	9.5
Digits Backward	DGB	12.9	—	11.2	9.1	12.1	8.9	12.3	13.4	12.9	17.3	9.5
Dots	DOT	14.8	14.7	—	10.7	13.2	10.5	13.5	14.4	14.0	18.2	10.9
Numbers & Letters Part A Speed	N&L _A -spd	12.1	12.0	14.0	—	11.6	8.2	11.8	12.9	12.4	17.0	8.9
Numbers & Letters Part A Errors	N&L _A -err	15.9	15.9	17.4	15.2	—	11.4	14.2	15.1	14.7	18.7	11.8
Numbers & Letters Part A Efficiency	N&L _A -eff	11.8	11.7	13.8	10.9	15.0	—	11.6	12.8	12.2	16.9	8.6
Numbers & Letters Part B Efficiency	N&L _B -eff	16.2	16.2	17.7	15.5	18.7	15.3	—	15.3	14.9	18.9	12.1
Numbers & Letters Part C Efficiency	N&L _C -eff	17.6	17.6	19.0	17.0	19.9	16.8	20.2	—	15.8	19.6	13.2
Numbers & Letters Part D Efficiency	N&L _D -eff	17.0	16.9	18.4	16.3	19.3	16.1	19.6	20.8	—	19.3	12.7
Numbers & Letters Part D Disruption	N&L _D -dis	22.9	22.8	24.0	22.4	24.7	22.2	24.9	25.8	25.4	—	17.2
Driving Scenes	DRV	12.6	12.5	14.4	11.7	15.6	11.3	15.9	17.3	16.7	22.6	—

Note. Differences required for statistical significance at the .05 level appear above the diagonal in the shaded area; differences significant at the .01 level appear below the diagonal. The differences required for statistical significance were calculated with the following formula: Difference score = $Z\sqrt{(SEM_a^2 + SEM_b^2)}$ where the Z is the normal curve value associated with the desired significance level and SEM_a and SEM_b are the standard errors of measurement of the scores.

Table 5.30
Discrepancies Between Language Module
Primary Scores Required for Statistical Significance
at the .05 and .01 Levels of Confidence for All Ages

Primary score	Acronym	OPD	AUD	NAM	WRT	BIL
Oral Production	OPD	—	17.0	14.8	16.3	16.1
Auditory Comprehension	AUD	22.4	—	16.8	18.1	17.9
Naming	NAM	19.5	22.1	—	16.1	15.9
Writing	WRT	21.5	23.9	21.2	—	17.3
Bill Payment	BIL	21.2	23.6	20.9	22.8	—

Note. Differences required for statistical significance at the .05 level appear above the diagonal in the shaded area; differences significant at the .01 level appear below the diagonal. The differences required for statistical significance were calculated with the following formula: Difference score = $Z\sqrt{(SEM_a^2 + SEM_b^2)}$ where the Z is the normal curve value associated with the desired significance level and SEM_a and SEM_b are the standard errors of measurement of the scores.

Table 5.31
Discrepancies Between Memory Module Primary Scores Required for
Statistical Significance at the .05 and .01 Levels of Confidence for All Ages

Primary score	Acronym	LLA-irc	LLB-irc	LLA-sd:drc	LLA-ld:drc	SHL-irg	SHL-drg	STL-irc:phu	STL-drc:phu	DLM-irc	DLM-drc
List Learning List A Immediate Recall	LLA-irc	—	15.7	13.7	14.2	13.5	13.3	16.0	17.0	11.7	13.2
List Learning List B Immediate Recall	LLB-irc	20.6	—	16.7	17.2	16.5	16.4	18.6	19.5	15.1	16.3
List Learning List A Short Delayed Recall	LLA-sd:drc	18.0	22.0	—	15.4	14.7	14.5	17.0	18.0	13.1	14.4
List Learning List A Long Delayed Recall	LLA-ld:drc	18.7	22.6	20.3	—	15.2	15.0	17.5	18.4	13.6	15.0
Shape Learning Immediate Recognition	SHL-irg	17.8	21.8	19.4	20.0	—	14.3	16.9	17.8	12.8	14.2
Shape Learning Delayed Recognition	SHL-drg	17.5	21.5	19.1	19.8	18.8	—	16.7	17.7	12.6	14.0
Story Learning Phrase Unit Immediate Recall	STL-irc:phu	21.1	24.5	22.4	23.0	22.2	22.0	—	19.8	15.4	16.6
Story Learning Phrase Unit Delayed Recall	STL-drc:phu	22.4	25.7	23.7	24.2	23.4	23.2	26.0	—	16.5	17.6
Daily Living Memory Immediate Recall	DLM-irc	15.4	19.9	17.2	17.9	16.9	16.6	20.3	21.7	—	12.5
Daily Living Memory Delayed Recall	DLM-drc	17.4	21.5	19.0	19.7	18.7	18.5	21.9	23.2	16.5	—

Note. Differences required for statistical significance at the .05 level appear above the diagonal in the shaded area; differences significant at the .01 level appear below the diagonal. The differences required for statistical significance were calculated with the following formula: Difference score = $Z\sqrt{(SEM_a^2 + SEM_b^2)}$ where the Z is the normal curve value associated with the desired significance level and SEM_a and SEM_b are the standard errors of measurement of the scores.

Table 5.32
Discrepancies Between Spatial Module Primary Scores Required for
Statistical Significance at the .05 and .01 Levels of Confidence for All Ages

Primary score	Acronym	VIS	DES	FGD-cpy	FGD-cpy:org	FGD-irc	MAP
Visual Discrimination	VIS	—	12.5	14.6	19.1	15.1	15.1
Design Construction	DES	16.5	—	11.0	16.5	11.6	11.7
Figure Drawing Copy	FGD-cpy	19.2	14.4	—	18.1	13.8	13.9
Figure Drawing Copy Organization	FGD-cpy:org	25.1	21.7	23.8	—	18.5	18.5
Figure Drawing Immediate Recall	FGD-irc	19.8	15.3	18.2	24.3	—	14.4
Map Reading	MAP	19.9	15.4	18.3	24.4	18.9	—

Note. Differences required for statistical significance at the .05 level appear above the diagonal in the shaded area; differences significant at the .01 level appear below the diagonal. The differences required for statistical significance were calculated with the following formula: Difference score = $Z\sqrt{(SEM_a^2 + SEM_b^2)}$ where the Z is the normal curve value associated with the desired significance level and SEM_a and SEM_b are the standard errors of measurement of the scores.

Table 5.33
Discrepancies Between Executive Functions Module
Primary Scores Required for Statistical Significance at
the .05 and .01 Levels of Confidence for All Ages

Primary score	Acronym	MAZ	JDG	CAT	WGN
Mazes	MAZ	—	13.2	8.0	10.5
Judgment	JDG	17.3	—	14.0	15.5
Categories	CAT	10.6	18.4	—	11.5
Word Generation	WGN	13.8	20.4	15.1	—

Note. Differences required for statistical significance at the .05 level appear above the diagonal in the shaded area; differences significant at the .01 level appear below the diagonal. The differences required for statistical significance were calculated with the following formula: Difference score = $Z\sqrt{(SEM_a^2 + SEM_b^2)}$ where the Z is the normal curve value associated with the desired significance level and SEM_a and SEM_b are the standard errors of measurement of the scores.

Table 5.34
Cumulative Percentages of the Demographically Corrected Standardization Sample
Obtaining Screening Domain/Index Score Discrepancies for All Ages Combined

		Screening Module Domain/Index pairs																	
Amount of discrepancy		% of sample earning domain/index pair score discrepancies																	
		S-ATT-S-LAN	S-ATT-S-MEM	S-ATT-S-SPT	S-ATT-S-EXE	S-ATT-S-NAB	S-LAN-S-MEM	S-LAN-S-SPT	S-LAN-S-EXE	S-LAN-S-NAB	S-MEM-S-SPT	S-MEM-S-EXE	S-MEM-S-NAB	S-SPT-S-EXE	S-SPT-S-NAB	S-EXE-S-NAB			
50		0.8	0.6	0.8	0.4		0.8	0.7	1.0	0.1	0.9	0.6		0.4					
49		0.9	0.7	1.0	0.5		1.0	0.9	1.3	0.1	1.0	0.7	0.1	0.4					
48		1.2	0.7	1.3	0.5		1.4	0.9	1.5	0.2	1.2	0.9	0.1	0.4					
47		1.5	0.8	1.3	0.6		1.5	1.2	1.7	0.3	1.4	1.1	0.1	0.5					
46		1.9	0.8	1.3	0.6		1.8	1.4	1.9	0.3	1.4	1.2	0.1	0.6	0.1				
45		2.1	1.1	1.4	0.8		2.0	1.7	2.0	0.3	1.7	1.4	0.2	0.7	0.3				
44		2.4	1.6	1.5	1.0		2.2	2.1	2.4	0.3	1.9	1.8	0.2	0.9	0.3				
43		2.9	1.8	2.0	1.0		2.8	2.2	2.6	0.4	2.3	2.0	0.2	1.0	0.3				
42		3.7	2.1	2.4	1.0		3.0	2.4	3.3	0.4	2.3	2.2	0.2	1.3	0.3				
41		4.0	2.2	2.7	1.2		3.6	2.8	3.7	0.6	2.6	2.8	0.3	1.6	0.3				
40		4.3	3.0	3.0	1.4		4.0	3.2	4.0	0.8	3.1	3.1	0.5	2.0	0.3				
39		4.9	3.3	3.5	1.7		4.7	3.5	4.7	0.8	3.9	3.2	0.6	2.4	0.3				
38		5.6	3.8	3.9	1.9		5.3	4.2	5.3	0.8	4.2	3.7	0.6	2.9	0.3				
37		6.0	4.4	4.5	2.2		6.0	4.7	5.7	1.0	4.5	4.2	0.7	3.3	0.3				
36		6.3	4.6	5.3	2.4	0.1	6.8	5.9	6.6	1.0	5.3	4.5	1.0	3.8	0.3				0.1
35		7.3	5.4	5.9	2.8	0.2	7.6	6.7	7.9	1.3	6.3	5.3	1.0	4.7	0.3				0.3
34		8.1	6.4	6.6	3.1	0.2	8.0	7.3	8.8	1.7	6.9	6.2	1.2	5.6	0.7				0.3
33		9.3	7.5	7.6	3.9	0.4	8.5	7.9	9.7	2.2	7.8	6.7	1.3	6.4	0.8				0.4
32		10.2	8.1	8.6	5.2	0.8	9.3	9.1	10.7	2.7	8.7	7.8	1.7	7.1	1.3				0.6
31		12.0	9.1	9.9	5.9	1.2	10.4	10.3	11.9	3.7	9.6	8.4	1.8	8.1	1.7				0.7
30		13.4	10.4	11.0	6.5	1.8	11.3	11.4	12.6	4.5	10.3	9.7	2.3	9.6	2.1				0.8
29		14.3	11.5	12.0	7.2	2.4	11.9	12.4	13.3	5.4	11.7	10.8	2.9	10.3	2.4				0.9
28		16.2	12.8	13.5	8.1	2.7	13.7	13.7	14.4	6.6	13.5	11.9	3.3	11.9	2.9				1.5
27		17.3	14.0	15.1	9.0	3.1	14.8	15.6	16.1	7.3	15.3	13.3	3.6	13.9	3.8				2.1
26		18.7	15.8	16.6	10.3	3.9	16.6	17.3	17.8	8.0	17.0	15.3	4.3	15.1	4.3				3.1
25		20.2	17.4	18.1	11.8	4.6	18.4	18.7	19.2	9.2	19.0	16.7	4.9	16.7	5.9				3.6
24		21.7	19.4	20.1	13.3	5.7	20.0	20.4	21.0	10.5	20.9	18.5	6.1	18.2	6.5				4.4
23		23.5	21.7	21.9	14.9	7.1	21.9	22.9	23.0	11.2	23.0	20.4	7.2	19.7	8.4				4.9
22		25.6	24.0	24.9	16.5	8.0	24.4	24.6	25.4	13.0	25.5	22.3	9.1	21.8	9.8				6.0

(continued)

Table 5.34 (continued)
Cumulative Percentages of the Demographically Corrected Standardization Sample
Obtaining Screening Domain/Index Score Discrepancies for All Ages Combined

Amount of discrepancy	Screening Module Domain/Index pairs															
	% of sample earning domain/index pair score discrepancies															
	S-ATT- S-LAN	S-ATT- S-MEM	S-ATT- S-SPT	S-ATT- S-EXE	S-ATT- S-NAB	S-LAN- S-MEM	S-LAN- S-SPT	S-LAN- S-EXE	S-LAN- S-NAB	S-MEM- S-SPT	S-MEM- S-EXE	S-MEM- S-NAB	S-SPT- S-EXE	S-SPT- S-NAB	S-EXE- S-NAB	S-EXE- S-NAB
21	27.1	26.8	27.2	18.8	9.4	26.5	27.0	27.5	15.7	28.2	25.5	10.7	24.5	11.8	7.2	7.2
20	30.4	29.2	30.1	20.6	11.4	28.6	30.0	30.0	17.8	30.1	27.0	12.5	26.3	13.1	8.4	8.4
19	33.1	31.1	32.4	23.1	13.5	31.4	32.5	32.3	20.3	32.1	29.6	14.2	28.5	15.3	10.2	10.2
18	35.6	33.6	34.6	26.3	15.7	33.6	35.5	35.3	22.6	34.4	32.2	17.5	31.3	17.6	12.6	12.6
17	38.9	37.1	37.1	28.4	18.1	36.0	38.2	38.0	25.1	36.7	35.5	20.4	34.1	20.2	15.3	15.3
16	41.9	39.7	39.8	31.3	20.8	39.4	41.7	40.8	27.5	39.8	38.1	23.8	36.8	22.7	18.3	18.3
15	44.8	42.2	43.1	34.5	23.8	43.4	44.7	44.4	31.0	42.9	41.4	27.3	39.5	26.0	22.1	22.1
14	48.3	45.4	46.9	37.6	26.4	47.0	48.0	47.9	34.0	45.9	44.5	31.7	43.2	29.7	25.8	25.8
13	51.5	48.3	50.6	41.0	29.6	50.5	51.5	51.5	36.4	49.5	47.8	34.6	46.5	33.1	29.7	29.7
12	54.8	51.5	54.3	45.6	33.8	54.0	55.3	55.3	41.0	52.9	51.3	37.5	50.2	37.3	33.9	33.9
11	58.4	55.0	58.0	49.7	38.8	58.1	58.0	59.7	45.6	56.5	55.8	41.5	54.6	41.8	37.8	37.8
10	62.6	58.1	62.2	54.0	43.9	62.3	62.2	62.8	49.3	60.2	58.8	46.5	58.0	46.9	41.5	41.5
9	65.6	62.2	65.7	58.0	48.1	66.8	65.5	65.5	53.6	63.5	61.6	51.1	61.5	52.1	45.4	45.4
8	69.9	66.5	68.9	62.0	53.0	71.3	68.4	70.0	58.1	67.4	65.7	55.9	66.6	56.7	50.2	50.2
7	74.0	70.4	72.5	67.0	58.4	74.8	72.6	73.8	63.5	71.5	70.0	61.5	71.3	62.2	55.9	55.9
6	77.2	75.3	76.0	73.2	64.0	78.3	76.6	78.3	69.2	74.7	75.3	66.6	75.4	67.3	62.2	62.2
5	80.3	79.5	80.7	76.8	69.6	82.2	80.6	83.5	74.3	80.1	80.8	72.4	80.2	72.5	68.5	68.5
4	84.6	83.1	84.4	81.9	75.2	86.5	85.0	87.4	80.0	84.5	85.1	78.3	84.2	77.5	74.7	74.7
3	88.8	88.5	89.5	87.0	81.7	89.9	88.9	91.0	85.9	89.4	90.1	85.1	88.6	84.4	81.2	81.2
2	93.2	93.5	93.8	92.9	89.4	93.7	93.2	94.2	91.0	93.3	93.9	90.8	93.4	90.2	89.1	89.1
1	98.2	97.7	98.1	97.6	96.4	97.7	97.8	97.4	96.9	97.6	97.6	96.8	97.6	96.1	97.0	97.0
0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
<i>M</i>	15.3	14.3	14.8	12.4	9.6	15.0	14.9	15.3	11.3	14.7	14.2	10.3	13.8	10.3	9.2	9.2
<i>Mdn</i>	13.0	12.0	13.0	10.0	8.0	13.0	13.0	13.0	9.0	12.0	12.0	9.0	12.0	9.0	8.0	8.0
<i>SD</i>	11.6	11.0	11.1	9.6	7.4	11.3	11.2	11.6	8.8	11.3	10.8	7.8	10.4	7.7	7.0	7.0

Note. *N* = 1,448. S-ATT = Screening Attention Domain; S-LAN = Screening Language Domain; S-MEM = Screening Memory Domain; S-SPT = Screening Spatial Domain; S-EXE = Screening Executive Functions Domain; S-NAB = Total Screening Index.

Table 5.35
Cumulative Percentages of the Age-Based, U.S. Census-Matched Standardization Sample
Obtaining Screening Domain/Index Score Discrepancies for All Ages Combined

		Screening Module Domain/Index pairs																	
		% of sample earning domain/index pair score discrepancies																	
Amount of discrepancy		S-ATT-S-LAN	S-ATT-S-MEM	S-ATT-S-SPT	S-ATT-S-EXE	S-ATT-S-NAB	S-LAN-S-MEM	S-LAN-S-SPT	S-LAN-S-EXE	S-LAN-S-NAB	S-MEM-S-SPT	S-MEM-S-EXE	S-MEM-S-NAB	S-SPT-S-EXE	S-SPT-S-NAB	S-EXE-S-NAB			
50		0.3	0.5	1.0	0.2		0.3	0.6	0.5		1.2	0.5				0.4			
49		0.3	0.6	1.0	0.2		0.4	0.6	0.5		1.3	0.7				0.5			
48		0.4	0.6	1.2	0.2		0.4	1.1	0.5		1.4	0.7				0.6			
47		0.6	0.6	1.3	0.3		0.4	1.5	0.5		1.4	0.8				0.6			
46		0.9	0.7	1.6	0.5		0.7	1.6	0.7		1.6	1.1				0.6			
45		1.2	0.9	1.6	0.6		1.1	1.6	1.0		1.8	1.1	0.1			0.6			
44		1.4	1.1	1.9	0.7		1.2	1.7	1.2		1.8	1.4	0.2			0.6			
43		1.5	1.3	2.0	0.7		1.4	2.1	1.6		2.1	1.6	0.4			0.9			
42		1.8	2.1	2.1	0.8		1.7	2.3	1.8		2.1	1.7	0.4		0.1	1.0			
41		2.0	2.1	2.3	0.8		2.1	2.8	2.1		2.4	1.8	0.5		0.2	1.4			
40		2.3	2.2	3.0	1.3		2.4	3.3	2.4		2.8	2.0	0.5		0.2	1.6			
39		2.7	2.4	3.3	1.4		2.5	3.4	2.6		3.4	2.3	0.5		0.2	1.7			
38		2.9	2.8	3.8	1.7		3.3	3.6	3.1		3.8	3.1	0.5		0.2	2.0			
37		3.8	3.2	4.4	2.1		3.6	4.6	3.6	0.4	4.3	3.4	0.9		0.3	2.3			
36		4.3	3.4	5.0	2.1		4.0	5.1	4.2	0.7	4.7	3.7	0.9		0.4	2.9			0.1
35		5.1	4.2	5.3	2.1	0.1	4.6	5.7	4.3	0.7	5.1	4.1	0.9		0.4	3.7			0.2
34		5.6	5.3	5.8	2.4	0.3	4.9	6.4	5.1	1.0	5.7	4.9	1.2		0.6	4.3			0.2
33		6.4	6.1	6.7	2.7	0.3	6.2	7.5	5.7	1.0	6.5	5.8	1.6		1.0	4.6			0.2
32		7.4	7.2	7.5	3.7	0.3	7.1	8.3	6.6	1.3	8.1	6.6	1.7		1.2	5.4			0.2
31		8.6	8.1	8.4	4.8	0.4	8.1	9.4	7.6	1.8	9.0	7.3	1.9		1.9	6.2			0.3
30		9.8	8.8	9.1	5.5	0.6	9.2	10.1	9.6	2.7	9.7	8.4	2.0		2.3	6.9			0.3
29		10.8	10.1	10.3	6.6	1.3	10.5	11.7	11.1	3.4	11.7	9.1	2.2		2.6	8.3			0.7
28		12.1	11.5	11.2	6.9	1.8	12.2	12.8	13.6	4.6	13.1	10.0	3.0		2.8	9.5			1.1
27		13.3	13.5	12.8	7.9	2.3	12.9	14.4	15.3	5.6	15.3	11.3	3.3		3.5	10.5			1.4
26		14.6	14.7	14.2	8.8	3.1	14.1	16.3	16.8	6.7	17.3	13.1	3.8		4.0	12.6			1.8
25		16.2	16.2	16.3	9.7	3.4	15.6	17.8	18.6	8.3	18.6	14.9	4.8		4.9	14.3			2.3
24		18.0	18.2	19.0	11.0	4.0	18.2	19.6	19.8	9.5	20.4	17.2	5.5		5.3	16.0			2.8
23		20.0	20.6	20.4	12.3	4.9	19.4	22.2	22.4	10.5	21.6	20.3	6.4		6.8	17.9			3.8
22		22.2	23.6	22.6	14.2	6.7	21.8	24.7	24.1	12.0	24.2	22.1	7.7		8.6	19.9			5.1

(continued)

Table 5.35 (continued)
Cumulative Percentages of the Age-Based, U.S. Census-Matched Standardization Sample
Obtaining Screening Domain/Index Score Discrepancies for All Ages Combined

Amount of discrepancy	Screening Module Domain/Index pairs															
	% of sample earning domain/index pair score discrepancies															
	S-ATT- S-LAN	S-ATT- S-MEM	S-ATT- S-SPT	S-ATT- S-EXE	S-ATT- S-NAB	S-LAN- S-MEM	S-LAN- S-SPT	S-LAN- S-EXE	S-LAN- S-NAB	S-MEM- S-SPT	S-MEM- S-EXE	S-MEM- S-NAB	S-SPT- S-EXE	S-SPT- S-NAB	S-EXE- S-NAB	
21	24.7	25.8	25.0	16.6	8.3	23.9	27.3	26.4	13.7	26.0	24.3	9.3	21.5	10.4	6.3	
20	26.9	28.0	27.0	18.9	9.8	26.9	29.7	28.5	16.0	28.4	25.8	11.0	24.0	12.3	7.6	
19	30.7	29.8	29.2	21.1	12.0	29.7	32.8	31.4	18.2	31.1	28.2	13.9	26.9	13.8	8.6	
18	33.3	33.4	32.3	24.2	13.5	34.2	35.2	34.1	21.1	34.1	30.5	15.9	29.1	16.5	10.0	
17	35.8	36.0	35.6	26.2	15.2	36.8	38.3	37.4	23.0	36.7	33.4	18.7	32.4	18.7	12.8	
16	38.7	38.0	38.8	28.9	17.8	39.7	41.2	41.5	25.4	40.1	36.7	21.7	35.2	21.4	15.0	
15	42.3	40.4	43.4	32.3	20.2	43.6	43.9	45.2	29.1	43.2	39.0	25.1	38.5	24.7	18.5	
14	45.3	43.6	46.4	35.2	23.0	47.3	47.2	49.0	32.0	46.2	42.1	27.7	41.6	28.5	21.3	
13	48.2	47.1	49.8	38.6	26.1	50.7	51.3	51.9	35.3	48.5	46.1	31.2	45.1	31.7	25.5	
12	51.3	50.2	53.5	43.6	29.4	54.2	55.4	54.7	39.6	51.9	49.6	34.4	46.8	34.0	30.0	
11	56.7	54.2	56.7	48.0	33.9	58.5	59.5	57.9	43.6	55.0	52.6	38.5	51.4	38.3	33.8	
10	60.4	56.6	61.0	52.0	38.5	62.2	62.7	62.2	48.2	59.3	57.2	43.6	56.1	43.0	38.5	
9	65.1	60.5	65.3	56.6	43.2	66.1	66.8	66.9	52.4	63.1	61.9	48.7	61.2	47.3	43.4	
8	68.9	65.8	69.3	61.4	49.6	68.9	69.1	71.4	57.9	67.6	65.3	53.5	64.3	53.5	48.1	
7	72.7	70.2	72.7	66.1	55.2	72.9	72.6	74.9	63.4	71.6	69.6	59.3	68.5	60.0	53.6	
6	76.5	75.4	77.2	70.9	61.9	76.8	79.2	78.4	67.7	75.7	73.8	65.9	73.9	67.0	59.8	
5	79.7	80.2	81.7	75.8	68.8	81.5	81.2	84.1	72.6	80.3	78.8	71.6	78.3	72.1	65.4	
4	85.3	84.4	85.2	82.2	75.2	85.2	84.2	87.2	78.9	84.9	83.0	77.8	83.4	78.3	71.6	
3	89.8	87.8	89.4	86.7	82.7	87.8	88.3	90.2	85.3	88.8	86.2	83.8	88.8	85.7	78.2	
2	93.7	92.3	93.1	93.7	89.9	92.9	91.9	94.6	91.2	91.9	91.2	90.9	92.9	91.3	87.1	
1	98.5	96.8	97.1	98.3	95.6	98.0	96.8	98.3	97.3	97.6	96.1	97.0	98.1	97.1	96.4	
0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	
<i>M</i>	14.2	13.9	14.4	11.9	9.0	14.3	14.8	14.7	10.8	14.5	13.5	9.9	13.2	9.9	8.5	
<i>Mdn</i>	12.0	12.0	12.0	10.0	7.0	13.0	13.0	13.0	9.0	12.0	11.0	8.0	11.0	8.0	7.0	
<i>SD</i>	10.5	10.6	10.7	9.1	6.9	10.3	11.0	10.4	8.1	11.1	10.5	7.6	9.9	7.5	6.6	

Note. *N* = 950. S-ATT = Screening Attention Domain; S-LAN = Screening Language Domain; S-MEM = Screening Memory Domain; S-SPT = Screening Spatial Domain; S-EXE = Screening Executive Functions Domain; S-NAB = Total Screening Index.

Table 5.36
Cumulative Percentages of the Demographically Corrected Standardization Sample
Obtaining Full Module and T-NAB Index Score Discrepancies for All Ages Combined

Amount of discrepancy	Module Index pairs														
	% of sample earning index pair score discrepancies														
	ATT- LAN	ATT- MEM	ATT- SPT	ATT- EXE	ATT- T-NAB	LAN- MEM	LAN- SPT	LAN- EXE	LAN- T-NAB	MEM- SPT	MEM- EXE	MEM- T-NAB	SPT- EXE	SPT- T-NAB	EXE- T-NAB
50		0.1	0.2	0.1		0.1	0.1	0.1		0.3	0.1				
49		0.2	0.3	0.1		0.1	0.3	0.2		0.3	0.1				
48		0.2	0.3	0.1		0.1	0.3	0.2		0.4	0.1				
47		0.2	0.3	0.2		0.2	0.3	0.3	0.1	0.5	0.1		0.1		
46	0.1	0.3	0.3	0.2		0.4	0.3	0.3	0.1	0.7	0.1		0.3		
45	0.2	0.5	0.3	0.3		0.5	0.3	0.3	0.1	0.8	0.3		0.3		
44	0.5	0.6	0.6	0.3		0.6	0.4	0.3	0.1	0.9	0.3		0.3		
43	0.5	0.6	0.8	0.3		0.8	0.5	0.3	0.1	1.0	0.5		0.5		
42	0.7	0.6	0.9	0.3		0.8	0.7	0.4	0.1	1.2	0.6		0.6		
41	0.7	0.6	1.0	0.3		0.9	0.7	0.6	0.1	1.4	0.8		0.8		
40	0.8	0.7	1.0	0.5		1.2	1.0	0.8	0.1	1.5	1.0		0.8		
39	0.9	0.8	1.2	0.6		1.3	1.3	1.2	0.1	1.8	1.2		1.1		
38	1.1	1.0	1.5	0.6		1.5	1.6	1.3	0.1	1.9	1.4		1.3		
37	1.5	1.1	1.9	0.6	0.1	1.7	1.9	1.6	0.2	2.1	1.7		1.4		
36	1.8	1.4	2.3	0.7	0.1	1.9	2.1	1.7	0.3	2.1	1.7		2.0		
35	2.0	1.6	2.6	1.0	0.1	2.3	2.4	2.1	0.3	2.9	1.8	0.1	2.4	0.1	
34	2.2	1.9	3.1	1.3	0.1	2.4	3.0	2.3	0.3	3.2	2.2	0.2	2.6	0.3	
33	3.1	2.3	4.0	1.6	0.1	3.1	4.0	2.8	0.3	3.5	3.0	0.3	3.3	0.4	0.1
32	3.9	3.2	4.3	1.9	0.1	3.7	4.7	3.1	0.3	4.2	3.5	0.4	3.8	0.6	0.2
31	4.5	3.7	4.6	2.2	0.3	4.1	5.8	3.3	0.3	4.8	4.1	0.4	4.3	0.8	0.2
30	5.5	4.2	5.1	2.4	0.3	4.9	6.2	4.3	0.6	5.3	4.9	0.7	4.8	0.8	0.2
29	6.6	5.2	5.8	3.1	0.3	5.8	7.2	5.1	0.7	5.7	5.6	0.9	5.6	1.0	0.3
28	7.4	6.7	6.4	3.8	0.6	6.7	8.0	6.0	1.0	6.9	6.3	1.0	6.4	1.3	0.3
27	8.4	7.7	7.1	4.6	0.7	7.4	9.2	7.0	1.4	7.6	7.0	1.1	7.2	1.6	0.3
26	9.8	9.0	8.4	5.9	0.9	8.1	10.1	8.4	1.6	8.3	8.1	1.3	8.6	2.0	0.5
25	11.2	10.2	9.7	6.7	1.3	9.3	11.2	9.5	2.3	10.0	9.0	1.7	10.0	2.3	0.7
24	13.1	11.8	10.7	8.2	1.4	11.2	12.6	10.4	2.9	11.3	11.3	2.3	11.4	2.7	1.2
23	14.9	13.9	12.0	10.2	1.7	13.0	14.5	12.3	3.9	13.1	12.9	2.7	12.8	3.6	1.5
22	16.9	15.5	13.5	11.3	2.4	14.5	17.0	13.9	5.0	14.3	14.3	3.4	14.2	4.3	1.9

(continued)

Table 5.36 (continued)
Cumulative Percentages of the Demographically Corrected Standardization Sample
Obtaining Full Module and T-NAB Index Score Discrepancies for All Ages Combined

Amount of discrepancy	Module Index pairs															
	% of sample earning index pair score discrepancies															
	ATT- LAN	ATT- MEM	ATT- SPT	ATT- EXE	ATT- T-NAB	LAN- MEM	LAN- SPT	LAN- EXE	LAN- T-NAB	MEM- SPT	MEM- EXE	MEM- T-NAB	SPT- EXE	SPT- T-NAB	EXE- T-NAB	
21	18.8	17.6	15.9	12.9	3.9	16.9	18.5	16.1	6.1	16.4	16.6	4.0	16.0	5.1	2.6	
20	20.9	19.6	17.9	14.5	5.0	19.8	21.0	18.4	6.8	18.8	18.8	5.4	17.8	6.1	3.8	
19	23.3	21.4	21.1	17.2	5.9	22.2	23.5	21.6	8.0	21.2	21.2	7.2	19.7	7.3	4.7	
18	25.9	23.7	24.6	20.3	7.3	25.0	26.1	23.9	9.4	23.5	23.6	8.9	22.2	8.7	6.0	
17	29.1	26.3	26.9	23.2	9.6	27.7	28.6	26.7	11.2	26.7	26.1	10.3	24.3	9.8	7.2	
16	32.0	29.4	29.6	25.5	11.4	30.1	30.5	29.5	13.2	29.3	28.7	12.8	27.8	11.7	9.2	
15	34.9	32.9	33.0	28.1	13.6	34.1	33.4	33.1	16.5	32.5	31.5	14.7	30.7	14.2	11.9	
14	39.7	36.0	36.7	31.9	16.6	37.5	36.4	36.4	19.4	36.7	35.0	17.1	34.3	17.4	15.2	
13	43.8	38.8	40.9	35.3	19.4	40.8	39.9	39.5	22.6	40.6	39.5	21.1	37.8	21.2	18.2	
12	47.5	42.5	44.3	38.2	24.1	44.3	44.6	42.5	27.1	44.8	43.4	24.2	41.3	25.2	21.6	
11	51.2	47.3	48.8	42.0	28.4	48.6	48.2	46.4	30.9	49.3	47.1	29.3	45.4	29.4	25.5	
10	55.2	51.0	53.2	46.8	33.4	52.9	52.2	49.4	35.4	53.3	51.2	34.5	50.7	35.5	30.1	
9	59.7	54.9	57.0	51.7	37.9	57.6	56.0	53.6	40.2	57.8	56.4	39.7	55.1	40.8	34.2	
8	64.8	59.8	61.4	56.5	44.5	62.4	60.2	58.8	46.1	62.3	60.7	46.8	60.1	47.4	39.9	
7	68.7	65.1	66.6	63.3	50.3	67.7	65.9	64.0	52.5	67.7	65.6	53.2	64.9	52.9	46.3	
6	73.3	70.6	71.2	69.1	57.2	72.1	71.7	68.9	58.0	72.1	69.3	59.8	69.8	59.9	53.7	
5	77.2	75.9	76.0	75.4	64.5	76.3	76.2	74.7	65.0	76.8	75.0	67.5	74.8	66.9	61.0	
4	82.1	81.0	81.4	79.6	72.6	81.2	81.0	79.9	72.5	81.7	80.8	74.4	80.9	73.5	70.6	
3	87.1	86.9	87.0	85.6	79.5	86.2	87.7	86.0	80.7	86.0	86.4	83.4	86.9	81.1	79.6	
2	92.3	92.4	92.5	90.5	87.4	91.8	92.4	91.3	87.8	91.5	91.4	88.8	92.2	89.2	87.6	
1	97.1	97.2	97.2	96.7	95.5	97.2	96.9	97.4	95.7	96.3	97.3	95.9	97.1	96.6	96.3	
0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	
<i>M</i>	12.4	11.8	11.9	10.7	7.8	12.0	12.2	11.6	8.3	12.0	11.7	8.2	11.6	8.2	7.3	
<i>Mdn</i>	11.0	10.0	10.0	9.0	7.0	10.0	10.0	9.0	7.0	10.0	10.0	7.0	10.0	7.0	6.0	
<i>SD</i>	9.1	9.0	9.1	8.2	5.9	9.1	9.4	9.0	6.5	9.3	9.0	6.1	9.0	6.3	5.6	

Note. *N* = 1,448. ATT = Attention Index; LAN = Language Index; MEM = Memory Index; SPT = Spatial Index; EXE = Executive Functions Index; T-NAB = Total NAB Index.

Table 5.37
Cumulative Percentages of the Age-Based, U.S. Census-Matched Sample Standardization Obtaining
Full Module and T-NAB Index Score Discrepancies for All Ages Combined

Amount of discrepancy	Module Index pairs														
	% of sample earning index pair score discrepancies														
	ATT- LAN	ATT- MEM	ATT- SPT	ATT- EXE	ATT- T-NAB	LAN- MEM	LAN- SPT	LAN- EXE	LAN- T-NAB	MEM- SPT	MEM- EXE	MEM- T-NAB	SPT- EXE	SPT- T-NAB	EXE- T-NAB
50			0.2							0.2					
49		0.1	0.2					0.1		0.2					
48		0.3	0.2			0.1	0.1	0.2		0.2					
47	0.1	0.3	0.3			0.2	0.2	0.2		0.3					
46	0.2	0.4	0.3			0.3	0.2	0.2		0.3	0.1		0.1		
45	0.2	0.5	0.3			0.3	0.3	0.2		0.3	0.2		0.2		
44	0.2	0.7	0.3	0.1		0.4	0.3	0.2		0.3	0.3		0.2		
43	0.2	0.7	0.3	0.3		0.6	0.3	0.2		0.3	0.3		0.2		
42	0.2	0.7	0.3	0.4		0.7	0.6	0.3		0.3	0.4		0.2		
41	0.3	0.7	0.3	0.4		0.8	0.8	0.5		0.5	0.6		0.2		
40	0.7	0.9	0.6	0.4		0.9	1.0	0.5	0.1	0.7	0.7		0.3		
39	0.8	1.1	0.6	0.4		1.1	1.2	0.7	0.2	1.2	0.7		0.5		
38	1.3	1.3	0.9	0.5		1.3	1.5	1.0	0.2	1.4	0.8		0.5		
37	1.4	1.4	1.1	0.6	0.1	1.4	1.7	1.3	0.3	1.6	1.0		0.7		
36	1.5	1.4	1.3	0.7	0.2	1.8	2.1	1.6	0.3	1.7	1.2		1.1		
35	1.6	1.6	1.4	0.7	0.2	2.3	2.2	1.8	0.3	1.7	1.4		1.3		
34	1.9	1.8	1.8	0.8	0.2	2.3	2.6	1.9	0.3	1.8	1.9	0.1	1.5		
33	2.0	1.9	2.3	0.8	0.4	3.2	3.1	2.1	0.3	2.0	2.4	0.2	1.6	0.1	
32	2.4	2.1	2.5	0.8	0.4	3.8	3.7	2.3	0.3	2.7	2.8	0.3	2.0	0.2	
31	3.2	2.6	2.7	1.3	0.4	4.0	4.4	2.5	0.3	3.1	3.4	0.3	2.4	0.2	0.1
30	3.7	3.1	3.2	1.6	0.4	4.5	5.0	2.9	0.4	3.4	4.0	0.3	3.2	0.3	0.2
29	4.3	4.2	3.7	2.1	0.4	5.4	5.6	3.5	0.5	4.3	4.3	0.4	3.9	0.4	0.2
28	5.7	5.7	4.0	2.8	0.4	5.5	6.4	4.0	1.0	5.3	4.9	0.5	4.7	0.4	0.2
27	6.3	6.8	5.1	3.2	0.4	6.2	7.6	5.0	1.0	6.2	5.9	1.0	5.6	0.7	0.3
26	7.6	8.2	6.1	3.9	0.4	6.6	8.9	5.9	1.3	7.6	6.9	1.1	6.6	1.0	0.5
25	8.7	8.9	7.1	5.4	0.7	7.9	9.7	6.9	1.7	8.6	8.3	1.4	7.6	1.2	0.5
24	10.1	10.8	8.3	6.1	0.9	9.7	10.8	8.2	2.0	10.1	9.4	2.0	8.7	1.5	0.5
23	11.6	12.4	9.4	7.5	1.4	10.9	12.5	9.7	2.7	11.9	10.7	2.6	10.0	1.7	0.5
22	13.3	14.0	10.9	9.1	1.7	12.2	14.9	11.9	3.2	13.3	12.7	3.3	11.9	2.6	1.1
															(continued)

(continued)

Table 5.37 (continued)
Cumulative Percentages of the Age-Based, U.S. Census-Matched Sample Standardization Obtaining Full Module and T-NAB Index Score Discrepancies for All Ages Combined

Module Index pairs															
Amount of discrepancy	% of sample earning index pair score discrepancies														
	ATT-LAN	ATT-MEM	ATT-SPT	ATT-EXE	ATT-T-NAB	LAN-MEM	LAN-SPT	LAN-EXE	LAN-T-NAB	MEM-SPT	MEM-EXE	MEM-T-NAB	SPT-EXE	SPT-T-NAB	EXE-T-NAB
21	15.0	16.2	12.9	10.5	2.3	13.4	17.0	13.3	5.0	15.7	14.9	4.0	13.3	3.5	1.6
20	17.0	18.3	14.3	12.4	2.9	14.9	18.6	15.3	5.4	17.9	16.3	4.8	14.6	4.5	2.1
19	19.9	19.8	16.6	13.8	3.9	17.5	21.1	17.6	6.7	19.3	17.8	6.2	16.7	5.5	2.8
18	22.3	23.1	19.3	16.3	5.2	21.0	23.3	20.1	8.2	21.5	19.5	8.4	18.5	6.6	4.4
17	25.3	25.1	22.9	19.1	7.1	23.6	25.4	23.5	9.4	24.8	21.5	9.0	21.5	8.0	6.0
16	28.5	28.1	26.2	22.7	9.1	26.6	28.8	26.4	10.8	27.6	24.7	10.6	24.9	10.2	6.8
15	32.2	31.1	29.3	25.4	10.8	29.9	30.9	29.2	12.6	30.0	28.4	12.8	28.6	11.8	9.1
14	35.9	33.7	32.7	30.0	13.7	33.1	35.0	32.8	16.4	33.1	32.2	15.3	32.5	13.4	12.2
13	40.4	37.8	36.4	32.7	17.0	37.1	37.4	36.2	19.6	37.8	35.8	18.4	36.1	16.7	14.7
12	43.3	41.3	41.1	35.8	21.5	40.7	41.2	39.9	23.3	42.3	40.4	23.0	38.6	20.4	18.5
11	47.3	44.8	45.7	39.6	25.6	45.0	44.5	42.6	27.8	45.6	44.8	26.0	42.5	25.5	21.9
10	52.5	48.8	50.6	45.0	29.8	49.4	51.2	46.8	33.4	49.5	48.4	30.9	45.8	30.5	26.4
9	56.6	53.9	53.6	50.2	35.7	54.1	55.6	52.1	37.9	53.0	52.6	36.6	51.3	35.5	32.3
8	62.5	58.9	59.2	55.2	41.4	58.9	60.6	57.6	42.6	58.1	57.0	41.4	56.6	41.9	36.1
7	67.9	63.2	64.7	62.0	48.7	65.1	65.5	63.3	48.8	63.6	62.4	48.8	63.0	47.7	42.8
6	72.7	67.2	70.1	68.1	55.0	70.6	69.8	68.4	55.7	68.1	69.1	56.0	69.2	54.4	50.7
5	76.9	72.1	74.2	73.5	62.6	74.6	75.6	74.7	63.5	73.6	75.0	61.5	73.5	63.3	59.1
4	81.2	78.3	80.8	80.5	68.9	79.2	79.8	79.4	70.5	79.2	80.0	69.4	80.0	71.4	67.5
3	86.4	84.0	85.7	85.4	78.2	85.1	85.6	84.4	80.0	86.8	85.8	79.5	86.1	79.8	76.3
2	91.3	89.8	92.3	90.7	87.1	91.4	91.4	89.7	87.1	91.0	91.6	87.8	91.4	88.5	86.4
1	96.8	95.8	97.7	97.5	94.4	97.2	97.5	96.7	96.0	97.4	97.7	96.1	97.0	96.6	94.9
0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
M	11.6	11.3	11.0	10.2	7.3	11.2	11.6	10.9	7.8	11.3	11.0	7.6	10.8	7.5	6.8
Mdn	10.0	9.0	10.0	9.0	6.0	9.0	10.0	9.0	6.0	9.0	9.0	6.0	9.0	6.0	6.0
SD	8.5	8.9	8.2	7.6	5.5	8.7	9.0	8.4	6.1	8.8	8.5	6.0	8.2	5.7	5.2

Note. *N* = 950. ATT = Attention Index; LAN = Language Index; MEM = Memory Index; SPT = Spatial Index; EXE = Executive Functions Index; T-NAB = Total NAB Index.