Extending Neuropsychological Assessments into the Primary Mental Ability Space

NO QUERY

Aging, Neuropsychology, and Cognition, 12:1–33

ISSN: 1382-5585/05
DOI: 10.1080/13825580590969343

Psychology Press
Taylor & Francis Group

Extending Neuropsychological Assessments into the Primary Mental Ability Space

K. WARNER SCHAIE[1], GRACE I. L. CASKIE[1], ANDREW J. REVELL[1], SHERRY L. WILLIS[1], ALFRED W. KASZNIAK[2] AND LINDA TERI[3]
[1]The Pennsylvania State University, [2]University of Arizona, and [3]University of Washington

ABSTRACT

A battery of 17 neuropsychological tests (including the CERAD battery) and 17 psychometric ability tests were administered to a sample of 499 participants of the Seattle Longitudinal study who had been given the psychometric ability tests seven and 14 years earlier. The neuropsychological tests were projected into a 5-factor psychometric ability space by means of extension analysis. The concurrent regressions of the neuropsychology tests on the psychometric ability tests were then used to estimate neuropsychology test scores from the psychometric ability tests administered in 1984, 1991 and 1998. Neuropsychologists then rated the study participants as either normal, suspect or cognitively impaired in 1998. Changes in estimated test scores were computed over seven and fourteen years. Significant odds ratios between normal and cognitively impaired groups were found for all neuropsychological tests over the proximal period and for most tests over the 14-year period. Similar findings occurred for the odds ratios between the normal and suspect groups for the most proximal 7-year changes.

INTRODUCTION

Most members of the psychological aging community utilize psychometric approaches to the measurement of cognitive status, cognitive change across age, and for the detection of cognitive deficits. However, the specific measurement systems that are utilized differ markedly depending on whether the investigators' interest is focused on the study of normal aging or the detection and diagnostic definition of neuropathology.

Manuscript Submitted to *Aging, Neuropsychology*
Original Submission: February 4, 2003
Revised: August 25, 2003

In order to study normal aging, it has generally been found necessary to construct assessment batteries that are suitable for measurement across the entire adult life span; hence requiring stimulus material across a wide range of difficulty. Measures typically used for this purpose are derived from L. L. Thurstone's (1938) work on defining primary mental abilities for the detailed study of normal intelligence (e.g., Ekstrom, French, Harman, & Derman, 1976; Horn, 1982; Schaie, 1985), or from the various forms of the Wechsler-Bellevue scales and its derivatives (Kaufman, Kaufman, McLean, & Reynolds, 1991; Matarazzo, 1972). A major characteristic of these approaches is that the measures are normally distributed in the population and that their factorial invariance over time can be demonstrated (e.g., Meredith, 1993; Schaie, Mailtland, Willis, & Intrieri, 1998). This allows measurement at the latent construct level.

By contrast, measures used by neuropsychologists (with the exception of the Wechsler scales whose use overlap both camps) are typically designed to have relatively low ceilings and bottoms because they are used to chart deficit from the point in time when it first noticed until the endpoint of death or total inability to respond to psychological measures is attained. A neuropsychological battery commonly used for the diagnosis of dementia was developed by the consortium to establish a registry for Alzheimer's disease (CERAD; Morris et al., 1989, 1993). Such measures usually do have non-normal highly skewed distributions, they are rarely measured at the latent construct level, and their factorial structure may differ for normal and abnormal population. Their purpose is not to identify the position of an examinee within the normal population, but rather to distinguish between normal and abnormal levels of function as denoted by cut-points. Since the objective of these measures is to detect the presence of pathology, they are not useful directly for the early detection of cognitive impairment or the identification of individuals at risk for the eventual occurrence of dementia.

Measures of normal aging, moreover, are generally designed to catalog a wide array of domains in order to permit profiling differential life courses and account for the maximum of individual differences in intellectual competence. By contrast, neuropsychological measures are usually targeted to detect specific diagnostic entities involved in the cognitive deficit, such as the diagnosis of Alzheimer's disease or deficits in executive functioning that may be associated with frontal lobe neuropathology. Hence, there has been only limited work on deriving latent construct for neuropsychological measures.

In this article, we therefore refer to measures that are designed to detect the presence of cognitive impairment as "neuropsychological" measures, and reserve the term "psychometric" measures for tests and constructs that are designed to assess an examinee's position in the general population.

The literature on early detection of risk for dementia in old age is fraught with the controversy whether or not early decline on any cognitive function might be a precursor of dementia or whether some declines simply represent age-related non-pathological changes in performance level that could be due

to disuse or other non-pathological factors (cf. Golden & Chronopolous, 1998; Vinters, 2001; Woodruff-Pak & Papka, 1999). Nevertheless, it is likely that excess cognitive decline may well be a most useful indicator of potential risk. There is also a strong likelihood that pharmaceutical and/or psychological interventions for the prevention of dementia will become available in the proximal future. As a consequence there is an increasing interest in the early detection of those at risk for cognitive impairment in old age so that interventions can be properly targeted (Bondi et al., 1995).

We are faced with the dilemma that the assessment procedures used by neuropsychologists for clinical diagnosis are not suitable for early detection of risk in their commonly used form, while psychometric measures useful for describing cognitive status in normal populations are not directly linked to the measures useful for clinical diagnosis. To overcome this problem there is a need to explore whether other psychometric measurement systems might be profitably employed for early detection. This would require efforts to project one measurement system into the measurement space of the other in order to develop suitable prediction equations.

We are reporting here the results of a study in which normal community-dwelling participants were administered an extensive primary mental abilities battery as well as an expanded CERAD battery including many measures commonly used in the clinical diagnosis of dementia. To determine the variance common to both batteries, it was necessary to identify a sample of normal individuals that included individuals who had begun to experience age-related cognitive decline and who would therefore show sufficient heterogeneity on the neuropsychological measures to permit cross-battery analyses. Such a sample is available in the latest cycle of the Seattle Longitudinal Study (SLS; Schaie, 1996, 2004). This sample has the further advantage that longitudinal psychometric data are available formany of the participants over time ranges from 7 to 42 years.

The purpose of the study then is to determine the projection of a neuropsychological battery developed for the detection of dementia into the normal mental ability factor space and to develop regression equations that permit post-diction of indicators of possible risk of dementia by considering study participants' longitudinal psychometric data at an age when neuropsychological assessment would not have been feasible or productive. By means of suitable transformations we create estimates of the neuropsychological measures that have psychometric properties (normality and invariance) that make them more appropriate for the assessment of individuals without clinical symptoms of dementia and avoid the need for cutpoints that would be unsuitable for a normal population. We then assess the effectiveness of utilizing the longitudinal psychometric data and the longitudinal change on the estimated neuropsychological data against the criterion of ratings made on the basis of a research protocol in a neuropsychological case conference.

METHOD

Study Participants

The sample consisted of 499 adults (211 men and 288 women) who were part of the SLS 7th wave data collection in 1997–98 and who ranged in age from 60 to 97 years (M = 73.07, SD = 8.30) at the time of their neuropsychological assessment. For the age/cohort group comparisons, we sub-divided the sample into an early-old group (age range 60–69 years, n = 180; m = 73, f = 107; M = 64.23, SD = 3.54), a middle-old group (age range 70–79, n = 205 m = 90, f = 115; M = 74.61, SD = 2.85), and an old-old group (age range 80–95, n = 114; m = 48, f = 66; M = 84.26, SD = 3.76). Educational level of the sample ranged from 7 to 20 years (M = 15.04, SD = 2.77). Participants were included in the neuropsychology studies only if they had been tested on the primary mental abilities battery on at least one previous occasion (7 years earlier).

The SLS is a longitudinal sequential study that was begun in 1956 and has assessed random samples from a large HMO in the Pacific Northwest. Participants are followed in 7-year intervals, and new random samples are recruited from the HMO population frame at each test occasions (see Schaie, 1996, for greater detail). All participants were community-dwelling and in average to excellent health for their age.

Assessment Procedure

The primary mental ability measures were administered to small groups of participants as part of a broader 5-hour battery conducted in two sessions each with breaks. Testing was conducted by an examiner with the assistance of a proctor. Testing locations were at familiar sites close to the homes of the participants. The neuropsychology battery was administered within one month of the mental ability battery. Neuropsychological testing was conducted individually in the participants' homes during a 2 1/2 hour session with a 15 minutes break.

Measures

The Primary Mental Abilities Battery

The SLS psychometric ability battery includes multiple measures marking each of six psychometric ability factors. A brief description of the primary abilities and the measures marking them is given below:

Inductive Reasoning

This ability involves identification of novel relationships in serial patterns and the inference of principles and rules in order to determine additional serial patterns.

PMA Reasoning (Thurstone & Thurstone, 1949). The participant is shown a series of letters (e.g., a b c b a d e f e) and is asked to identify the next letter in the series.

ADEPT Letter Series (Blieszner, Willis, & Baltes, 1981). This is a parallel form to the PMA Reasoning test.

Word Series (Schaie, 1985). The participant is shown a series of words (e.g., January, March, May) and is asked to identify the next word in the series. Positional patterns used in this test are identical to the PMA Reasoning test.

ETS Number Series (Ekstrom et al., 1976). The participant is shown a series of numbers (e.g., 6, 11, 15, 18, 20) and is asked to identify the next number that would continue the series.

Spatial Orientation

This is the ability to visualize and mentally manipulate spatial configurations, to maintain orientation with respect to spatial objects, and to perceive relationships among objects in space.

PMA Space (Thurstone & Thurstone, 1949). The study participant is shown an abstract figure and is asked to identify which six other drawing represents the model in two-dimensional space.

Object Rotation (Schaie, 1985). The participant is shown a line drawing of a meaningful object (e.g., an umbrella) and is asked to identify which of six other drawings represent the model rotated in two-dimensional space.

Alphanumeric Rotation (Willis & Schaie, 1983). The participant is shown a letter or number and is asked to identify which six other drawings represent the model rotated in two-dimensional space.

Test stimuli in the Object and Alphanumeric Rotation tests have the same angle of rotation as the abstract figures in the PMA Space test.

Cube Comparisons. (Ekstrom, et al., 1976) In each item, two drawings of a cube are presented; the participant is asked to indicate whether the two drawings are of the same cube, rotated in three-dimensional space.

Numerical Facility

This is the ability to understand numerical relationships and compute simple arithmetic functions.

PMA Number (Thurstone & Thurstone, 1949). The participant checks whether additions of simple sums shown are correct or incorrect.

Addition (Ekstrom, et al., 1976). This is a test of speed and accuracy in adding three single or two-digit numbers.

Subtraction and Multiplication (Ekstrom, et al., 1976). This is a test of speed and accuracy with alternate rows of simple subtraction and multiplication problems.

Verbal Comprehension

Language knowledge and comprehension is measured by assessing the scope of a person's recognition vocabulary.

PMA Verbal Meaning (Thurstone & Thurstone, 1949). A four-choice synonym test which is highly speeded.

ETS Vocabulary II (Ekstrom, et al., 1976). A five-choice synonym test of moderate difficulty level.

ETS Vocabulary IV (Ekstrom, et al., 1976). Another five-choice synonym test consisting mainly of difficult items.

Perceptual Speed

This is the ability to find figures, make comparisons and carry out other simple tasks involving visual perception, with speed and accuracy.

Identical Pictures (Ekstrom, et al., 1976). The participant identifies which of five numbered shapes or pictures in a row are identical to the model at the left of the row.

Finding A's (Ekstrom, et al., 1976). In each column of 40 words, the participant must identify the five words containing the letter "a".

Number Comparison (Ekstrom, et al., 1976). The participant inspects pairs of multi-digit numbers and indicates whether the two numbers in each pair are the same or different.

Verbal Recall

This is the ability to encode, store and recall meaningful language units.

Immediate Recall (Zelinski, Gilewski, & Schaie, 1993). Participants study a list of 20 words for 3 1/2 minutes. They are then given an equal period of time to recall the words in any order.

Delayed Recall (Zelinski et al., 1993). Participants are asked to recall the same list of words as in Immediate Recall after an hour of intervening activities (other psychometric tests).

PMA Word Fluency (Thurstone & Thurstone, 1949). The participant freely recalls as many words as possible according to a lexical rule within a five-minute period.

All tests are slightly speeded to be suitable for group administration. The longitudinal markers included in this battery (i.e., the original PMA tests, Thurstone & Thurstone 1949), by necessity (i.e., for consistency across

successive test administrations), employ the test booklet and answer sheet format used since the beginning of the SLS (Thurstone & Thurstone, 1949). However, print size on answer sheets has been enlarged from the original. All other forms use disposable booklets with enlarged type upon which answers are marked directly (cf. Ekstrom, French, Harman, & Derman, 1976; Schaie, 1985).

The Neuropsychological Battery

This battery consists of the CERAD measures (Morris et al., 1989, 1993), selected tests from the WAIS-R and the WMS-R and some other commonly used neuropsychological assessment instruments.

The *Mini-Mental State Examination* (MMSE; Crum, Anthony, Bassett, & Folstein, 1993; Folstein, Folstein, & McHugh, 1975; Lemsky, Smith, Malec, & Ivnik, 1996; Tombaugh & McIntyre, 1992; Uhlmann & Larson, 1991) is a cognitive screening test to assess participants' orientation to time and place, short term and delayed memory recall, ability to follow simple directions, praxis, and language. Although its psychometric characteristics have been questioned, we include this test to link with existent literature and to obtain a better understanding of how this screening instrument projects into the domains commonly measured in the assessment of older normal community-dwelling populations.

Verbal Fluency: "Animal Category" (Borkowski, Benton, & Spreen, 1967; Isaacs & Kennis, 1973; Welsh et al., 1994). The test measures impairment in verbal production, semantic memory and language, and is sensitive to the early changes in dementing illnesses such as Alzheimer's disease (Cohn et al., 1995). Participants name as many items as they can within 60 seconds in a given category

Modified Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1984; Morris et al., 1989; Van Gorp, Satz, Kiersch, & Henry, 1986) involves the verbal identification of two-dimensional objects within 20 seconds. If the name is not produced in that time, a semantic cue is given, and after another 20 seconds a phonemic cue is provided if the participant has difficulty. It is used clinically to measure impairment of language functions, and is sensitive to the early phases of progressive dementia (Cohn et al., 1995) The abbreviated version of this test is suitable for our population and requires no more than fifteen minutes.

Word List Memory Recall: Immediate and Delayed (Atkinson & Shiffrin, 1971; Cahn et al., 1995). Participants are presented with a list of 10 words and asked to recall as many words as they can in three trials and after a delay.

Constructional Praxis (Rosen, Mohs, & Davis, 1984; Welsh et al., 1994. The test, designed to assess persons with Alzheimer's disease, involves four line drawings in increasing complexity. Immediate and delayed recall were measured, but only delayed recall was included in our analyses.

Word List Recognition (Cahn et al., 1995; Mohs, Kim, & Johns, 1984). Participants must recognize and identify the words from the Word List

Memory task when presented among 10 distractor words. Sensitivity to mild dementia has been demonstrated (Cahn et al., 1995).

Sub-tests from the Wechsler Adult Intelligence Scale (WAIS-R). A short form of the WAIS-R (Wechsler, 1981) was also given. The short form consists of the most commonly used tests from both Verbal and Performance scales; tests that show early as well as late decline in old age.

Vocabulary test. This vocabulary test is the most commonly used measure of maintained verbal functions in clinical practice and clinically oriented research.

Comprehension test. A measure of common knowledge, may reflect intactness of logical thought.

Digit symbol substitution test. A speeded measure involving the matching of symbols and numbers.

Block design test. This is the classical clinical test of spatial visualization and has sometimes been used by neuropsychologists to identify problems in the visuo-motor pathways.

Digit Span test. A measure of short-term memory requiring the recall of forward and backward number series.

The Wechsler Memory Scale-Revised (WMS-R; Cahn et al., 1994; Ryan, Paolo, & Brugardt, 1990, 1992; Wechsler, 1981) is one of the oldest clinical instruments for assessing memory impairment. However, we only use Logical memory, with immediate and delayed recall because other parts of the WMS overlap with the other measures in our battery.

The Trail-Making Test is one of the earliest measures used by neuropsychologists to detect difficulty in attention and cognitive inflexibility (Cahn et al., 1994; Heaton, Grant, & Matthews, 1986; Reitan & wolfson, 1985). Part A requires tracing a consecutively ordered path among a set of numbers. Part B involves tracing a path that requires shifting between numbers and letters. The Trail-Making Test is sensitive to the early phases of dementing illnesses such as Alzheimer's disease (Lafleche & Albert, 1995), and to change to dementia progression (Storandt, Botwinick, Danziger, Berg, & Hughes, 1984).

The Fuld Object Memory Test (Fuld, 1977) is a free recall of objects measure. We included the sub-scales for retrieval and rapid verbal retrieval in our analyses.

The Mattis Dementia Rating Scale (MDRS; Mattis, 1988; Vitaliano, Russo, Breen, Vitello, & Prinz, 1986) is a sensitive index of cognitive functioning in dementia patients. It yields a total score and five sub-scale scores (attention, initiation and perseveration, construction, conceptualization, and memory). It provides good discrimination between normal and cognitively impaired groups (Green, Woodard, & Green, 1995; Monsch et al., 1995; Vangel & Lichtenberg, 1995). The MDRS has also been found be sensitive to change as dementia progresses (Kiyak, Teri, & Borson, 1994; LaRue, 1992; Smith et al., 1994).

The neuropsychological battery included the following additional measures which were not included in the analyses reported in this article: the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977; Radloff & Teri, 1986), the McMaster Problem Solving scale (Epstein, Baldwin, & Bishop, 1983; Epstein, Bishop, Ryan, Miller, & Keitner, 1993); the instrumental activities of daily living (IADL; Lawton & Brody, 1969); and a metamemory measure, the Memory Functioning Questinnaire (MFQ; Gilewski & Zelinski, 1988). These measures are not primarily measures of cognition and therefore would not be expected to project substantially into the primary mental abilities domains.

Neuropsychologist Ratings

Given the nature of our community dwelling samples we did not have any medical examinations or clinical dementia ratings. Instead we relied on a research protocol involving a two-step procedure for the rating of participants' neuropsychological functional status. First, participants' were evaluated against a screening algorithm to determine whether a given record had characteristics that might result in a rating of cognitive impairment in a neuropsychological case conference. The screening algorithm utilized cutoff scores that were selected based upon previous research indicating a positive association between meeting the cutoff criteria and cognitive dysfunction (Crum, Anthony, Bassett, & Folstein, 1993; LaRue, 1992; Spreen & Strauss, 1991). The cutoff criteria for the selected tests are:

1. MMSE-score <27
2. Mattis Dementia Rating Scale-score < 130
3. Trail B-score time > 180 seconds
4. An age adjusted scaled score < 7 for any of the following: WAIS-R Vocabulary, WAIS-R Comprehension, WAIS-R Block Design, and WAIS-R Digit Symbol.

As a second step, those records that met the algorithm's screening criteria were then examined in detail by two neuropsychology consultants. In the consensus conferences, scores on the neuropsychological tests, and tester's report of observed sensory limitations and current or previous health problems were considered. The likelihood of decline from a previous level was also evaluated by considering the participant's education and occupation as well as the presence of decline on psychometric tests over the previous seven years. Participants received one of the following ratings: 1) the participant is normal, 2) the participant does not have evidence of dementia at this time but has one or more characteristic that suggests further monitoring is indicated, 3) the participant probably has evidence of dementia, or 4) the participant definitely has evidence of dementia.

The neuropsychological ratings identified 354 participants (70.9%) as normal, 111 participants (22.2%) to require monitoring, 22 participants (4.4%) to have probable evidence of dementia, and 12 (2.4%) to have definite evidence of dementia. There were no significant gender differences in the proportions of individuals assigned to the different rating classifications. As was to be expected there were significant age differences between rating groups. The group requiring monitoring was approximately 4 years older than the normal group, and the categories with dementia were 8 years older than the normal group. There were no educational differences between the normal group and the groups with evidence of dementia, but the "monitor" group had a approximately one year less education on average than both the normal group and the groups with dementia. Mean CES-D scores for the four groups were 7.26, 8.92, 11.41, and 12.02 respectively. Reported mean IADL complaints were 0.83, 1.01, 2.05, and 3.00.

Statistical Procedures

Analysis Plan

The data analysis plan involved first the confirmation of the factor structure for the primary mental ability measures. Second, an extension analysis was conducted to determine the relation of the neuropsychology measures to the primary mental abilities. Third, the regressions of the primary mental abilities on the neuropsychology measures were used to estimate neuropsychology measures for prior SLS occasions. Fourth, change scores for the primary mental abilities and the estimated neuropsychology scores were computed from 1984 to 1991 and from 1991 to 1998; and participants were classified as to whether they had experienced reliable decline or not.

Transformations

For ease of comparisons all raw data were transformed to T-scores with a mean of 50 and a standard deviation of 10. Neuropsychological variables with skewness greater than 2.00 were normalized using a McCall transformation (Garrett, 1966). The normalized variables were: Fuld Retrieval, the MMSE, Word List Recognition, and Trails A. Also, values above 300 seconds on Trails B were trimmed to a value of 300 before T-score transformation.

RESULTS

Descriptive Data

Descriptive data for the variables included in this study are provided in Tables 1 and 2. Because information on a community dwelling sample on this extensive data base may beof broader interest, we are reporting means and standard deviations by gender and age/cohort group as well as for the total sample. Table 3 presents the intercorrelations

TABLE 1. Raw Score Means and Standard Deviations for the Primary Mental Abilities Battery

	Early Old Age			Middle Old Age			Old-Old Age			Total	Total	Total
Variable	Males	Females	Total	Males	Females	Total	Males	Females	Total	Males	Females	All
Reasoning												
PMA	17.46	18.08	17.83	13.98	15.10	14.61	9.70	11.91	10.99	14.23	15.48	14.95
Reasoning	(5.43)	(5.62)	(5.54)	(5.86)	(5.18)	(5.50)	(5.08)	(5.74)	(5.56)	(6.23)	(5.94)	(6.09)
ADEPT	11.30	10.75	10.97	8.90	9.02	8.97	6.47	7.48	7.06	9.19	9.31	9.26
Letter series	(3.89)	(3.60)	(3.72)	(3.69)	(3.02)	(3.32)	(3.63)	(3.35)	(3.49)	(4.14)	(3.54)	(3.80)
Word Series	18.32	18.60	18.48	14.78	16.09	15.51	10.89	13.09	12.18	15.04	16.33	15.83
	(5.43)	(5.22)	(5.29)	(5.23)	(4.35)	(4.79)	(5.39)	(5.25)	(5.40)	(5.99)	(5.31)	(5.63)
Number Series	7.61	5.79	6.49	5.80	5.16	5.44	4.28	4.43	4.37	6.08	5.23	5.59
	(2.82)	(2.88)	(2.97)	(2.92)	(2.68)	(2.80)	(2.58)	(2.54)	(2.55)	(3.05)	(2.76)	(2.92)
Spatial Orientation												
PMA Space	26.12	20.11	22.55	19.32	16.73	17.87	11.85	12.52	12.24	19.98	17.02	18.27
	(10.96)	(10.43)	(11.02)	(9.22)	(8.81)	(9.06)	(8.86)	(8.34)	(8.54)	(11.10)	(9.75)	(10.43)
Object Rotation	40.33	35.42	37.41	33.76	30.18	31.75	23.45	22.61	22.96	33.73	30.39	31.80
	(11.46)	(11.45)	(11.67)	(11.49)	(11.80)	(11.77)	(14.40)	(12.25)	(13.14)	(13.64)	(12.69)	(13.19)
Alphan. Rotation	41.68	39.43	40.34	33.34	35.20	34.89	25.70	27.79	26.92	34.54	35.07	34.85
	(10.10)	(11.55)	(11.01)	(12.98)	(11.77)	(12.31)	(12.07)	(11.27)	(11.60)	(13.23)	(12.34)	(12.71)
Cube	22.86	17.46	19.65	17.44	15.81	16.52	14.51	13.21	13.76	18.67	15.85	17.05
Comparison	(5.51)	(5.44)	(6.18)	(5.86)	(5.47)	(5.69)	(5.44)	(5.43)	(5.44)	(6.59)	(5.66)	(6.22)
Verbal Ability												
PMA Verbal	40.92	41.10	41.03	36.61	38.27	37.54	27.96	32.29	30.46	36.13	37.95	37.18
Meaning	(8.07)	(9.02)	(8.62)	(10.08)	(9.36)	(9.24)	(11.86)	(10.34)	(11.16)	(10.96)	(10.01)	(10.45)
ETS Vocabulary	30.00	30.07	30.04	29.53	30.48	30.06	28.57	30.00	29.41	29.48	30.22	29.91
	(3.93)	(3.93)	(3.92)	(5.04)	(4.18)	(4.59)	(5.92)	(5.71)	(5.81)	(4.91)	(4.48)	(4.68)
ETS Adv.	26.47	25.63	25.97	24.98	27.04	26.14	24.74	26.32	25.66	25.44	26.35	25.97
Vocabulary	(6.09)	(5.75)	(5.89)	(5.56)	(6.02)	(5.90)	(7.61)	(6.97)	(7.25)	(6.27)	(6.16)	(6.22)
Numeric Ability												
PMA Number	24.11	23.07	23.49	23.80	21.70	22.62	16.94	20.21	18.83	22.34	21.87	22.07
	(10.46)	(9.80)	(10.12)	(10.64)	(7.92)	(9.24)	(9.31)	(9.11)	(9.30)	(10.66)	(9.00)	(9.73)

TABLE 1. Continued

Variable	Early Old Age			Middle Old Age			Old-Old Age			Total	Total	Total
	Males	Females	Total	Males	Females	Total	Males	Females	Total	Males	Females	All
ETS Addition	42.89	41.74	42.21	40.66	40.88	40.78	35.74	38.18	37.17	40.33	40.58	40.48
	(14.10)	(14.47)	(14.29)	(13.53)	(12.28)	(12.81)	(12.67)	(12.77)	(12.73)	(13.74)	(13.26)	(13.45)
Subtaction &	55.67	51.25	53.04	50.04	48.95	49.43	40.45	43.74	42.37	49.85	48.61	49.13
Multiplication	(19.82)	(19.27)	(19.58)	(16.56)	(17.43)	(17.02)	(18.07)	(16.80)	(17.34)	(18.87)	(18.17)	(18.46)
Perceptual Speed												
Identical Pictures	35.36	33.25	34.11	27.61	28.17	27.92	21.74	24.36	23.27	28.99	29.18	29.10
	(7.03)	(5.75)	(6.37)	(6.69)	(5.58)	(6.08)	(5.86)	(6.22)	(6.18)	(8.29)	(6.73)	(7.47)
Number	22.34	23.39	22.97	18.41	21.16	19.95	15.89	19.29	17.88	19.21	21.56	20.57
Comparison	(5.07)	(5.03)	(5.06)	(4.29)	(4.83)	(4.79)	(3.83)	(5.04)	(4.86)	(5.11)	(5.18)	(5.28)
Finding A's	26.75	29.38	28.32	23.37	26.92	25.36	22.09	25.85	24.28	24.26	27.59	26.18
	(7.09)	(9.69)	(9.09)	(6.85)	(7.88)	(7.63)	(6.99)	(7.56)	(7.53)	(7.49)	(8.62)	(8.32)
Verbal Memory												
PMA Word	41.47	43.51	42.68	35.49	38.34	37.09	29.75	34.95	32.76	36.25	39.49	38.12
Fluency	(12.67)	(12.83)	(12.77)	(12.02)	(12.69)	(12.45)	(12.33)	(12.07)	(12.40)	(13.03)	(13.00)	(13.10)
Immediate Recall	12.88	14.74	13.98	11.11	12.98	12.17	9.62	10.89	10.36	11.39	13.16	12.41
	(3.66)	(3.60)	(3.73)	(3.83)	(4.00)	(4.02)	(4.09)	(3.94)	(4.04)	(4.01)	(4.09)	(4.15)
Delayed Recall	10.56	13.35	12.22	8.47	11.16	9.99	7.11	8.52	7.93	8.89	11.37	10.33
	(4.53)	(4.43)	(4.67)	(4.36)	(4.52)	(4.64)	(4.54)	(4.57)	(4.59)	(4.64)	(4.84)	(4.91)

TABLE 2. Raw Score Means and Standard Deviations for the Neuropsychology Battery

	Early Old Age			Middle Old Age			Old-Old Age			Total		
Variable	Males	Females	Total	Males	Females	Total	Males	Females	Total	Males	Females	All
CERAD												
Boston Naming	14.64	14.53	14.58	14.30	14.30	14.30	13.83	13.41	13.59	14.31	14.18	14.24
	(0.63)	(0.70)	(0.68)	(1.01)	(0.91)	(0.95)	(1.52)	(1.48)	(1.50)	(1.09)	(1.09)	(6.71)
MMSE	28.62	28.87	28.77	27.82	28.50	28.20	27.40	27.61	27.52	28.00	28.43	28.25
	(1.46)	(1.43)	(1.45)	(2.94)	(1.65)	(2.33)	(2.56)	(2.37)	(2.44)	(2.47)	(1.83)	(2.13)
Praxis (Delayed)	8.76	8.49	8.60	8.08	7.51	7.76	6.08	6.02	6.04	7.86	7.54	7.67
	(1.98)	(2.28)	(2.16)	(2.23)	(2.73)	(2.53)	(3.50)	(2.93)	(3.17)	(2.69)	(2.78)	(2.74)
Verbal Fluency	21.47	21.19	21.30	18.11	18.67	18.42	17.17	16.61	16.84	19.06	19.13	19.10
	(5.83)	(5.03)	(5.35)	(4.90)	(4.84)	(4.86)	(6.16)	(4.85)	(5.42)	(5.79)	(5.21)	(5.45)
Word List	7.62	8.47	8.12	6.49	7.37	6.99	5.65	6.66	6.23	6.69	7.62	7.23
‘ (Recall)	(1.81)	(1.79)	(1.84)	(2.10)	(1.95)	(2.06)	(2.50)	(2.42)	(2.49)	(2.22)	(2.13)	(2.21)
WAIS-R												
Digit Span	16.01	16.23	16.14	15.53	14.66	15.04	13.92	14.83	14.45	15.33	15.28	15.30
	(3.69)	(4.31)	(4.06)	(4.48)	(3.40)	(3.93)	(3.89)	(3.90)	(3.91)	(4.15)	(3.93)	(4.02)
Vocabulary	58.15	54.57	56.02	53.84	55.19	54.60	50.67	53.67	52.40	54.61	54.61	54.61
	(6.98)	(8.52)	(8.10)	(8.83)	(9.03)	(8.95)	(13.11)	(9.18)	(11.05)	(9.82)	(8.86)	(9.27)
Comprehension	24.42	22.43	23.24	23.42	22.90	23.12	21.54	22.29	21.97	23.34	22.58	22.90
	(3.73)	(4.61)	(4.37)	(4.50)	(4.52)	(4.51)	(5.00)	(4.80)	(4.87)	(4.48)	(4.61)	(4.57)
Block Design	32.14	29.37	30.49	27.08	25.01	25.91	19.60	19.85	19.75	27.13	25.45	26.16
	(8.30)	(8.91)	(8.75)	(8.78)	(7.90)	(8.34)	(8.07)	(7.88)	(7.92)	(9.62)	(9.01)	(9.30)
Digit Symbol	49.85	50.51	50.24	41.57	42.29	41.98	32.02	35.80	34.19	42.27	43.88	43.20
	(9.65)	(10.46)	(10.12)	(8.65)	(9.90)	(9.36)	(10.77)	(9.75)	(10.32)	(11.58)	(11.54)	(11.57)
OTHER NEUROPSYCHOLOGY MEASURES												
WMS-R	28.82	29.15	29.02	24.31	26.12	25.33	22.52	23.82	23.33	25.47	26.74	26.21
Immediate	(6.64)	(6.90)	(6.78)	(7.69)	(6.80)	(7.24)	(8.60)	(7.90)	(8.20)	(7.95)	(7.36)	(7.63)
WMS-R	24.66	23.97	24.25	18.21	20.83	19.68	16.17	17.79	17.11	19.99	21.30	20.74
Delayed	(8.31)	(7.78)	(7.99)	(8.39)	(7.94)	(8.22)	(9.87)	(9.05)	(9.40)	(9.36)	(8.46)	(8.87)

TABLE 2. Continued

	Early Old Age			Middle Old Age			Old-Old Age			Total		
Variable	Males	Females	Total	Males	Females	Total	Males	Females	Total	Males	Females	All
Trails A	35.32	35.48	35.41	45.39	42.08	43.52	61.90	54.35	57.56	45.66	42.40	43.78
	(15.63)	(15.42)	(15.46)	(17.82)	(14.28)	(15.97)	(40.21)	(22.24)	(31.23)	(26.07)	(18.18)	(21.90)
Trails B	80.00	87.76	84.61	109.05	106.43	107.56	158.23	147.72	152.17	109.97	108.68	109.22
	(29.36)	(46.70)	(40.65)	(52.83)	(45.61)	(48.76)	(71.12)	(69.12)	(69.84)	(58.68)	(56.64)	(57.45)
Fuld Retrieval	44.34	46.21	45.45	42.38	43.33	42.92	37.46	39.34	38.82	41.94	43.63	42.91
	(4.02)	(2.71)	(3.42)	(4.96)	(4.47)	(4.70)	(10.41)	(8.35)	(9.32)	(6.85)	(5.65)	(6.24)
Fuld Rapid Verbal Retrieval	65.42	74.10	70.58	59.21	66.75	63.46	51.79	60.20	56.60	59.68	68.04	64.50
	(10.58)	(13.08)	(12.83)	(11.24)	(11.62)	(12.02)	(15.83)	(11.64)	(14.16)	(13.19)	(13.26)	(13.85)
Mattis Total	139.27	139.55	139.44	136.19	138.83	137.67	131.58	136.42	134.39	136.21	138.55	137.56
	(4.51)	(3.99)	(4.24)	(12.01)	(4.39)	(8.68)	(10.63)	(5.38)	(8.33)	(10.09)	(4.64)	(7.53)

TABLE 3. Intercorrelations Among the Psychometric and Neuropsychology Measures (N = 499)

Measure	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Correlations among the Neuropsychology Measures																
1. PMA Reasoning	.78	.81	.63	.56	.54	.54	.52	.65	.35	.38	.47	.44	.50	.48	.49	.27
2. ADEPT Letter Series		.75	.64	.51	.48	.49	.51	.56	.33	.36	.46	.40	.46	.42	.47	.45
3. Word Series			.60	.51	.51	.51	.49	.61	.37	.38	.44	.41	.48	.48	.49	.47
4. Number Series				.46	.41	.43	.48	.53	.27	.30	.46	.48	.51	.35	.33	.34
5. PMA Space					.77	.66	.55	.46	.22	.22	.35	.33	.33	.27	.25	.23
6. Object Rotation						.70	.51	.49	.24	.19	.32	.30	.32	.27	.26	.24
7. Alphanumeric Rotation							.52	.50	.25	.20	.32	.32	.36	.29	.32	.32
8. Cube Comparison								.42	.10	.13	.33	.32	.38	.28	.23	.21
9. PMA Verbal Meaning									.52	.53	.47	.48	.52	.50	.51	.47
10. ETS Vocabulary II										.79	.23	.21	.25	.37	.35	.27
11. ETS Vocabulary IV											.23	.21	.25	.46	.39	.32
12. PMA Number												.82	.72	.35	.26	.26
13. ETS Addition													.81	.38	.28	.28
14. Subtraction/Multiplication														.38	.33	.33
15. PMA Word Fluency															.46	.43
16. Immediate Recall																.90
17. Delayed Recall																

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Correlations between the Neuropsychology and Psychometric Measures																	
18. Boston Naming	.27	.26	.29	.24	.30	.31	.27	.26	.34	.27	.24	.11	.06	.12	.19	.23	.23
19. FULD Retrieval	.41	.39	.44	.24	.28	.33	.30	.21	.42	.20	.20	.22	.23	.26	.32	.54	.57
20. FULD Rapid Retr.	.46	.39	.50	.29	.26	.27	.34	.24	.51	.31	.35	.29	.34	.36	.59	.50	.50
21. Mattis Total	.44	.42	.44	.30	.35	.33	.30	.20	.44	.33	.37	.30	.28	.32	.39	.47	.44
22. WMS-R Immediate	.42	.44	.44	.28	.27	.26	.23	.19	.40	.35	.39	.16	.14	.17	.29	.46	.45
23. WMS-R Delayed	.42	.46	.41	.31	.28	.29	.24	.22	.40	.32	.38	.18	.16	.22	.33	.53	.54
24. WAIS-R Digit Span	.37	.34	.41	.34	.25	.20	.20	.15	.34	.33	.35	.35	.20	.26	.40	.26	.25
25. WAIS-R Vocabulary	.42	.38	.40	.31	.28	.26	.26	.20	.50	.65	.73	.23	.24	.25	.45	.41	.35

TABLE 4. Continued

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Correlations between the Neuropsychology and Psychometric Measures																	
26. WAIS-R Compreh.	.30	.30	.32	.27	.26	.22	.20	.16	.37	.44	.51	.14	.15	.19	.29	.32	.29
27. WAIS-R Block Des.	.57	.56	.53	.53	.56	.51	.52	.52	.50	.27	.27	.34	.33	.36	.33	.37	.35
28. WAIS-R Digit Sym.	.62	.53	.59	.52	.49	.48	.50	.51	.62	.19	.19	.43	.51	.58	.42	.45	.27
29. MMSE	.40	.45	.45	.35	.26	.23	.26	.22	.44	.33	.33	.39	.33	.38	.36	.47	.44
30. Verbal Fluency	.36	.36	.33	.29	.25	.28	.25	.24	.43	.36	.41	.14	.18	.25	.37	.40	.34
31. Word List Recall	.31	.32	.30	.17	.20	.19	.19	.14	.37	.18	.24	.14	.16	.21	.30	.59	.60
32. Praxis Delayed	.40	.41	.40	.36	.38	.36	.36	.31	.39	.19	.21	.20	.19	.26	.30	.38	.38
33. Trails A	.46	.39	.46	.35	.41	.40	.41	.40	.44	.13	.12	.27	.30	.34	.36	.34	.33
34. Trails B	.59	.55	.59	.50	.46	.50	.48	.46	.59	.25	.24	.42	.44	.47	.40	.42	.40

Measure		19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
Correlations among the Neuropsychology Measures																	
18. Boston Naming		.26	.29	.29	.31	.28	.13	.32	.30	.33	.27	.24	.32	.22	.32	.28	.34
19. FULD Retrieval			.54	.44	.46	.51	.21	.27	.25	.37	.50	.42	.35	.62	.47	.39	.48
20. FULD Rapid Retrieval				.45	.41	.45	.29	.43	.33	.41	.54	.40	.51	.44	.37	.45	.47
21. FULD Rapid Retrieval					.41	.40	.31	.50	.46	.36	.43	.41	.34	.39	.36	.30	.40
22. WMS-R Immediate						.90	.21	.45	.39	.32	.31	.42	.40	.53	.40	.24	.35
23. WMS-R Delayed							.17	.44	.37	.33	.34	.44	.42	.59	.43	.25	.36
24. WAIS-R Digit Span								.36	.30	.29	.28	.32	.17	.14	.14	.24	.30
25. WAIS-R Vocabulary									.69	.39	.26	.34	.44	.29	.33	.21	.32
26. WAIS-R Comprehension										.37	.23	.29	.36	.29	.31	.21	.26
27. WAIS-R Block Design											.55	.38	.38	.28	.46	.46	.52
28. WAIS-R Digit Symbol												.37	.38	.39	.41	.60	.62
29. MMSE													.31	.40	.33	.27	.40
30. Verbal Fluency														.35	.31	.28	.36
31. Word List Recall															.44	.24	.32
32. Praxis Delayed																.30	.42
33. Trails A																	.57
34. Trails B																	

All correlations are significant at the .05 level of confidence.

among the 17 primary mental ability measures and the 17 neuropsychology measures.

Primary Mental Abilities Factor Structure

The fit of the six-factor structure for the 20 primary mental ability tests employed in the SLS (Schaie, Dutta, & Willis, 1991) was assessed for the present sample. All factor models were estimated using the full information maximum likelihood procedure implemented in Amos 4.0 (Arbuckle & Wothke, 1999), This procedure estimates the model parameters from the raw data matrix, rather than from a covariance or correlation matrix.

When the 17 neuropsychology measures were added to this battery, it was found that a permissible 6-factor solution could not be obtained because of colinearity of many of the neuropsychology measures with the perceptual speed factor. We have previously shown that substantial proportions of individual differences in speeded tests are absorbed by perceptual speed (Schaie, 1989). This effect becomes even more problematic in older samples because of increasing convergence of abilities (sometimes referred to as dedifferentiation; cf. Schaie, 2000). We determined therefore hat it would be necessary to remove the Perceptual Speed factor and the related observed measures from the abilities battery in order to achieve optimal estimation of the neuropsychology measures in the proposed extension analysis.

The factor structure for the PMA battery minus the three perceptual speed tests was recomputed for the remaining 17 variables and five factors based on the sample used in the present study. The fit for the reduced five factor solution was X^2 (df = 108, N = 499) = 536.08, $p < .001$, CFI = .99, RMSEA = .09, TLI = .98. Standardized factor loadings were significant for all salient values reported by Schaie et al. (1991). Hence, the five mental ability factors included in the extension analysis were Inductive Reasoning, Spatial Orientation, Numerical Facility, Verbal Comprehension, and Verbal Memory.

Extension Analyses for the Neuropsychology Measures

An important application of confirmatory factor analysis is to use this procedure to implement the Dwyer (1937) extension method. As Tucker (1971) demonstrated, it is not appropriate to use factor scores on a latent variable to estimate their regression on an observed variable. However, confirmatory factor analysis permits the estimation of the location of some new observed variable or variables of interest within a previously known factor (latent construct) space. This is a situation that frequently arises in aging studies as samples are followed over long time periods.

To conduct an optimal extension analysis it is necessary to have a sample for whom data are concurrently available both on a set of measures whose dimensionality (i.e. latent constructs) have been well established as well as the other measures whose relation to these constructs is to be studied.

For our purposes, we began with the psychometric abilities battery that has been employed in the SLS since 1983. We then added the CERAD as well as other neuropsychological measures that we wished to relate to the psychometric ability dimensions.

In the extension analysis, factor loadings were constrained to the unstandardized values from the confirmatory factor analysis solution for the cognitive variables for this sample. Factor loadings for the neuropsychological measures were then freely estimated providing information on the projection of these measures into the previously established five-factor cognitive factor structure. Because multiple scores from several of the neuropsychology tests were used, three residual covariances were estimated. Trails A with Trails B, Fuld Retrieval with Fuld Rapid Verbal Retrieval, and WMS-R Immediate with WMS-R Delayed. Factor variances for the five latent cognitive factors were fixed to unity. Error variances for the 34 observed variables were freely estimated.

As might be expected, the neuropsychological assessment measures, when extended into the psychometric abilities factor structure, generally spread over two or more of the psychometric ability domains (see Table 4). All measures, except for the WAIS-R Digit Span, Vocabulary, Comprehension and Block Design scales, had significant loadings on the Verbal Memory factor. Of the latter scales, Digit Span, Vocabulary and Comprehension

TABLE 4. Standardized Loadings of Neuropsychological Tests on Cognitive Factors Allowing Correlated Errors for Sub-tests of Fuld, WMS-R, and Trails

Neuropsychology Test	Reasoning	Spatial	Verbal	Number	Memory
Boston Naming Test – CERAD	.11	.41***	.29***	.16**	.11*
FULD Retrieval[a]	.05	.28***	.01	.02	.58***
FULD Rapid Verbal Retrieval	.00	.16**	.25***	.09*	.37***
Mattis Grand Total[a,b]	.04	.21***	.34***	.03	.29***
WMSR Immediate Total	.13	.11	.33***	.19***	.33***
WMSR Delayed Total	.02	.17**	.26***	.15**	.47***
WAISR Digit Span	.30***	.10	.31***	.11*	.05
WAISR Vocabulary	.16**	.13**	.93***	.08*	.04
WAISR Comprehension	.22**	.22***	.74***	.11*	.08
WAISR Block Design	.17*	.50***	.13**	.02	.06
WAISR Digit Symbol	.19**	.37***	.11**	.25***	.23***
MMSE[a]	.13	.00	.19***	.15**	.30***
Verbal Fluency	.10	.24***	.41***	.07	.24***
Word List Recall	.20**	.20***	.07	.06	.71***
Praxis Delayed Total	.03	.42***	.10*	.09	.31***
Trails A[a]	.10	.41***	.07	.07	.18***
Trails B[c]	.23**	.34***	.02	.15***	.17***

* $p < .05$; ** $p < .01$; *** $p < .001$

[a] Normalized with McCall transformation, [b] Extreme low values trimmed to 102; [c] Extreme high values trimmed to 300

had their largest extensions into the Verbal Comprehension factor, while Block Design extended most prominently into the Spatial Ability factor.

Most measures also had a secondary loading on the Spatial Ability factor, except for the Wechsler Memory Immediate Recall, the WAIS-R Digit Span scale, and the MMSE. Several measures also had secondary and/or tertiary loadings on the Inductive Reasoning and Numeric Ability factors. The negative loadings found for Trails were expected because, for that measure, a large score (time to completion) is in the unfavorable direction.

Regression of Neuropsychology Variables on Earlier PMA status

Using factor weights obtained by orthonormal transformation of the values in Table 4, we first estimated T-scores on the neuropsychology measures from the PMA factor scores for the concurrent occasion to obtain information on the relation between estimated and observed T-scores. Table 5 reports the correlations between the observed and estimated neuropsychology test scores as well as the multiple correlations between the concurrent PMA tests and the neuropsychology tests both with and without including age and education as predictors. As can be seen, the values from the extension analyses are somewhat more conservative because they attenuate for error of measurement.

We conclude that we can validly estimate scores on the neuropsychology tests from scores on the five PMA factors on the basis of the following

TABLE 5. Concurrent Prediction of Neuropsychology Tests from the Primary Mental Ability Factors, OLS Regression and Extension Analyses

Neuropsychology Test	Multiple *R* from OLS Regression	Multiple *R* from OLS Regression incl. age and education	Correlation of Estimated Scores from Extension Analysis
Boston Naming Test – CERAD	.406	.452	.363
FULD Retrieval	.594	.615	.584
FULD Rapid Verbal Retrieval	.595	.600	.574
Mattis Grand Total	.570	.511	.542
WMS-R Immediate Total	.573	.578	.526
WMS-R Delayed Total	.604	.611	.578
WAIS-R Digit Span	.483	.489	.461
WAIS-R Vocabulary	.752	.765	.746
WAIS-R Comprehension	.538	.562	.532
WAIS-R Block Design	.678	.703	.639
WAIS-R Digit Symbol	.731	.819	.708
MMSE	.580	.586	.550
Verbal Fluency	.519	.566	.499
Word List Recall	.612	.626	.607
Praxis Delayed Total	.516	.521	.501
Trails A	.558	.569	.512
Trails B	.680	.704	.664

All values are statistically significant, $p < .001$.

considerations: First, all correlations between estimated and observed neuropsychology scores are significant at the .001 confidence level. Second, the correlations between the observed and estimated correlations for the neuropsychology measures approach the reliable variance of the tests as reported in the measures section of this article. Third, the correlations between observed and estimated scores are also within the first decimal for alternate OLS regression estimates for most measures. However, the extension analysis derived estimated scores are to be preferred because they adjust for error of measurement (Tucker, 1971). Hence it seemed reasonable to attempt backwards prediction (post-diction) to estimate what our participants' earlier scores on the neuropsychology battery might have been if we had had the opportunity to measure them seven and fourteen years earlier.

We next used the factor weights from the extension analyses to estimate (post-dict) T-scores for the neuropsychology tests for our data collections that occurred seven (1991) and fourteen years (1984) prior to the direct measurement on the neuropsychology tests.

Decline in neuropsychology measures by age group

Mean values by age group (young-old, old-old, very-old) are provided for the three estimated data points in Table 6. Age declines significant at the .01 level of confidence are observed in the young-old group over a 14-year interval (1984–1998) for all measures except the Mattis scale. However, significant decline on the WAIS-R Vocabulary scale is observed only over the second 7-year interval from 1991 to 1998.

In the old-old group, significant change over 7 years from 1984 to 1991 is found for WAIS-R Digit Symbol, Praxis Delayed Total, Trails A and Trails B. Significant 14-year changes (1984–1998) occur for all measures except the Mattis scale. These can be attributed primarily due to the many significant declines occurring during the 1991 to 1998 period. Finally, in the very-old group, significant 7-year changes from 1984 to 1991 are found for the Boston Naming Test, WAIS-R Block Design and Digit Symbol scales as well as Praxis Delayed Total, Trails A and Trails B. Significant 14-year changes are found for all measures.

Predicting Dementia Ratings from Longitudinal Data

We next examined the relative effectiveness of utilizing longitudinal primary mental abilities data and the estimated neuropsychology data in predicting ratings made by our neuropsychologists. We first examine change over the most proximal seven years from 1991 to 1998. Then we reach back another seven years and examine changes occurring from 1984 to 1991. Longitudinal change is considered both for the PMA factor scores (computed from the actual observations) and for the estimated neuropsychology measures. In each instance we first contrast all participants rated as having

TABLE 6. Predicted Neuropsychology Test Score Means for 1984, 1991 and 1998 by Age Group

	Young-Old			Old-Old			Very Old		
Neuropsychology Test	1984	1991	1998	1984	1991	1998	1984	1991	1998
Boston Naming Test	60.03	59.59	55.49[ab]	56.99	55.74	50.73[ab]	52.50	49.91[a]	43.52[ab]
FULD Retrieval	58.22	59.84	54.83[ab]	55.53	54.57	49.70[ab]	53.27	51.31	44.63[ab]
FULD Rapid Verbal Retr.	57.36	58.76	54.18[ab]	55.42	54.57	49.94[ab]	53.11	51.43	44.96[ab]
Mattis Grand Total	53.19	54.16	54.31	51.72	51.09	50.52	49.25	47.67	45.02[ab]
WMS-R ImmediateTotal	56.94	58.03	53.80[ab]	55.63	54.81	50.30[ab]	53.08	51.49	45.06[ab]
WMS-R Delayed Total	57.41	58.84	54.21[ab]	55.46	54.60	49.84[ab]	53.24	51.54	45.04[ab]
WAIS-R Digit Span	57.67	57.72	54.20[ab]	55.52	55.16	50.52[ab]	51.18	50.21	44.86[ab]
WAIS-R Vocabulary	52.64	52.95	51.42[b]	53.09	53.27	50.98[ab]	51.36	51.09	47.44[ab]
WAIS-R Comprehension	53.82	54.21	52.19[ab]	53.78	53.78	50.93[ab]	51.55	50.94	46.78[ab]
WAIS-R Block Design	61.50	60.90	56.52[ab]	57.22	55.89	50.65[ab]	52.62	49.50[a]	43.31[ab]
WAIS-R Digit Symbol	61.59	61.63	56.31[ab]	58.22	56.53[a]	50.57[ab]	53.58	50.35[a]	43.12[ab]
MMSE	57.56	58.88	54.21[ab]	55.71	54.85	49.90[ab]	53.22	51.48	45.00[ab]
Verbal Fluency	56.52	57.14	53.77[ab]	55.24	54.78	50.70[ab]	52.43	51.02	45.24[ab]
Word List Recall	57.73	59.43	54.55[ab]	55.20	54.31	49.54[ab]	53.08	51.33	45.05[ab]
Praxis Delayed Total	61.29	61.70	56.35[ab]	57.50	55.95[a]	50.33[ab]	53.45	50.34[a]	43.09[ab]
Trails A	61.94	61.66	56.69[ab]	57.64	56.14[a]	50.53[ab]	53.08	50.00[a]	43.04[ab]
Trails B	62.07	61.88	56.72[ab]	57.87	56.31[a]	50.37[ab]	52.77	49.69[a]	42.71[ab]

All values are in T-score units scaled on the total 1998 sample (Mean = 50, SD = 10). [a] Significant decline in performance from baseline (1984 score) at 01 level of confidence. [b] Significant decline from 1991 performance.

some suspicious characteristics against the normal participants (Rating 1 vs. combined Ratings 2, 3 and 4). We then contrast only those individuals who were identified as having evidence of probable or definite dementia against the normal group (Ratings 1 vs Rating 3 and 4). In tables 7 through 10, we consequently distinguish between Normals (Rating 1), suspect (Ratings, 2, 3 and 4) and those with dementia (Ratings 3 and 4). Data are reported only for the total sample because there were no statistically significant sex x rating category or age group x rating category interactions. In each case we report mean longitudinal change in T-score points. Perhaps of greater practical interest, however, is our report of the proportion of individuals who show reliable decline (defined as a drop that is equal or greater than 1 SE from Time 1, as well as the odds ratios between the normal and diagnosed groups.

Changes over the most proximal seven years (1991–1998)

PMA factor scores. Table 7 provides average declines in T-score points, proportions of the rating groups who declined significantly over seven years, and the odds ratios of these proportions contrasting the normal and rating groups.

Given that all of the participants of this study are over sixty (mean age = 73 years at the time they were rated), it is not surprising that we observed significant average age changes on all of the factor scores. There is a significant interaction between magnitude of 7-year change and rating group for all factor scores except for Inductive Reasoning. As expected, greater change is observed for the groups rated as other than normal. When we contrast all individuals with some suspicious characteristics with normals, significant odds ratios are obtained only for the Verbal Comprehension and the Verbal Recall factors. However, when only those rated as having probable or definite dementia are contrasted with the normals, significant odds ratios are found for all estimated neuropsychology measures.

Estimated neuropsychology scores. Data for the estimated neuropsychology scores may be found in Table 8. Again, significant interactions are found between magnitude of 7-year change and rating groups, with greater change for both the suspect and dementia categories. Odds ratios are statistically significant for the suspect group for all neuropsychology scores except the Boston Naming Test and for Word List Recall. All odds ratios are significant for the group with dementia. It is noteworthy, that the odds ratios for the estimated neuropsychology measures are substantially larger than those for the psychometric factor scores.

Changes over the earlier seven year period (1984–1991)

Having established that we can provide meaningful estimates over the most proximal seven years (1991–1998) prior to the actual neuropsychological assessment of our study participants, we then reached further back to determine

TABLE 7. Mean Decline in T-Score Points, Proportion of Subjects Declining from 1991 to 1998 and Odds Ratios of Diagnosed vs Normal Groups for the PMA Factor Scores

	Mean Decline			Proportion Declining			Odds Ratios	
Factor	Normal	Suspect	With Dementia	Normal	Suspect	With Dementia	Suspect	With Dementia
Inductive Reasoning	3.77	4.41	6.24	53.9	59.6	72.7	1.10	1.35*
Spatial Orientation	3.47	4.41	5.95	40.9	47.2	58.8	1.15	1.44*
Numeric Facility	3.15	4.09	7.26	37.8	44.4	64.7	1.17	1.71**
Verbal Comprehension	1.42	2.55	6.03	21.3	37.3	61.8	1.75*	2.90***
Verbal Recall	4.13	6.39	10.57	34.5	46.5	64.7	1.30**	1.88***

* $p < .05$; ** $p < .01$; *** $p < .001$

TABLE 8. Mean Decline in T-Score Points, Proportion of Subjects Declining from 1991 to 1998 and Odds Ratios of Diagnosed vs Normal Groups for the Estimated Neuropsychology Scores

	Mean Decline			Proportion Declining			Odds Ratios	
Factor	Normal	Suspect	With Dementia	Normal	Suspect	With Dementia	Suspect	With Dementia
Boston Naming Test	4.26	5.84	8.67	50.3	55.9	78.6	1.11	1.40*
FULD Retrieval	5.23	7.81	13.13	37.3	50.4	66.7	1.35**	1.79***
FULD Rapid Verbal Retr.	4.21	6.44	11.23	40.5	57.4	75.8	1.48***	1.70***
Mattis Grand Total	0.80	2.44	5.53	13.1	34.0	57.6	2.60***	4.34***
WMS-R Immed. Total	3.42	6.41	10.14	43.3	59.6	87.9	1.36***	2.02***
WMS-R Delayed Total	4.48	7.28	12.00	41.6	55.3	75.8	1.33**	1.82***
WAIS-R Digit Span	2.31	5.05	6.88	41.6	52.5	60.6	1.26*	1.32*
WAIS-R Vocabulary	0.39	1.82	1.50	21.1	35.5	57.6	1.68***	2.73***
WAIS-R Comprehension	0.26	2.56	2.74	23.6	39.0	60.6	1.65***	2.56***
WAIS-R Block Design	5.06	5.83	8.73	47.3	58.9	66.7	1.24*	1.41*
WAIS-R Digit Symbol	5.73	7.64	12.21	57.0	73.0	90.6	1.28***	1.59***
MMSE	4.33	7.34	12.00	42.2	56.0	81.8	1.33**	1.94***
Verbal Fluency	2.26	4.80	6.82	37.6	55.3	75.8	2.67***	2.01***
Word List Recall	5.02	7.67	13.04	38.5	44.5	63.6	1.24	1.65**
Praxis Delayed Total	5.89	7.47	12.05	47.3	64.4	75.8	1.36***	1.60***
Trails A	5.42	6.74	10.47	41.6	58.2	69.7	1.40***	1.68**
Trails B	5.76	7.32	11.53	53.0	67.4	81.8	1.27**	1.54***

* $p < .05$; ** $p < .01$; *** $p < .001$

the effectiveness of this procedure in identifying individuals at risk at an earlier point in time by studying the predictive effectiveness of change over the preceding 7-year period.

PMA factor scores. Table 9 provides data on change on the PMA factor scores from 1984 to 1991 (the end point is now seven years prior to the actual administration of the neuropsychology tests. Participants at T1 = 1984 in this analysis were in their late fifties. Hence, decline over seven years was not significant for any PMA factor for the normal and suspect groups. However, significant odds ratios were found for the group with dementia ($p < .05$) for Numeric Facility and Verbal Comprehension.

Estimated neuropsychology scores. Results for the estimated neuropsychology scores are given in Table 10. Significant interactions between magnitude of 7-year decline and rating group were found for all measures except the Boston Naming Test, WAIS-R Digit Span, WAIS-R Block Design, Praxis Delayed total, and part A of the Trail-making Test. Significant odds ratios when contrasting the suspect with the normal group were found for the estimated scores of the Fuld Retrieval, the Delayed Wechsler Memory, the MMSE, and for Word List Recall. Significant odds ratios contrasting the group with dementia with the normal group were obtained for all measures except Boston Naming, WAIS-R Digit Span, WAIS-R Block Design, Praxis Delayed total, and part A of the Trail-making test.

DISCUSSION AND CONCLUSIONS

This study used the method of extension analysis to project a battery of neuropsychological measures into a five-factor primary mental ability factor structure. The location of the neuropsychological measures within the mental ability factor space was assessed via significant factor loadings of the neuropsychological tests on five mental ability factors. In order to obtain a permissible solution, it was necessary to remove the perceptual speed factor from the test battery used to locate the neuropsychology measures in the primary mental ability space. This action was necessitated by the colinearity of the perceptual speed measures with the neuropsychology measures. Past work has suggested, moreover, that substantial proportions of individual differences variance in speeded cognitive measures are absorbed by the perceptual speed factor (cf. Schaie, 1989). Moreover, sensory and central changes associated with normal aging lead to the convergence of the ability space in advanced age (cf. Baltes & Lindenberger, 1997; Schaie, 2000). Decline in speed of performance is widely acknowledged as a central phenomenon in the reduction of cognitive resources in normal aging (cf. Salthouse, 1999). However, it does not seem to be directly predictive of the occurrence of impairment in executive functions. Hence, our decision to exclude the perceptual speed measures was guided by both theoretical and empirical considerations.

TABLE 9. Mean Decline in T-Score Points, Proportion of Subjects Declining from 1984 to 1991 and Odds Ratios of Diagnosed vs Normal Groups for the PMA Factor Scores

	Mean Decline			Proportion Declining			Odds Ratios	
Factor	Normal	Suspect	With Dementia	Normal	Suspect	With Dementia	Suspect	With Dementia
Inductive Reasoning	0.50	0.44	1.81	20.0	19.0	31.0	0.95	1.54
Spatial Orientation	1.28	1.50	1.54	22.2	24.0	26.9	1.08	1.21
Numeric Facility	1.54	1.98	3.69	25.8	33.0	46.2	1.32	1.79*
Verbal Comprehension	0.21	0.12	1.42	9.8	12.4	23.1	1.26	2.35*
Verbal Recall	0.29	1.26	3.00	2.0	2.1	3.8	1.08	1.96

* $p < .05$; ** $p < .01$; *** $p < .001$

TABLE 10. Mean Decline in T-Score Points, Proportion of Subjects Declining from 1984 to 1991 and Odds Ratios of Diagnosed vs Normal Groups for the Estimated Neuropsychology Scores

	Mean Decline			Proportion Declining			Odds Ratios	
Factor	Normal	Suspect	With Dementia	Normal	Suspect	With Dementia	Suspect	With Dementia
Boston Naming Test	1.22	1.54	2.85	21.3	22.7	30.8	1.07	1.45
FULD Retrieval	0.01	1.39	2.96	16.5	33.0	38.5	1.99***	2.33***
FULD Rapid Verbal Retr.	0.02	1.21	3.27	12.6	22.7	42.3	1.80*	3.36***
Mattis Grand Total	0.07	0.93	2.81	20.1	28.9	53.8	1.44	2.68***
WMS-R Immediate Total	0.05	1.13	3.50	14.5	20.6	46.2	1.42	3.17***
WMS-R Delayed Total	0.06	1.31	3.35	13.4	30.9	46.2	2.31***	3.45***
WAIS-R Digit Span	0.30	0.28	2.50	16.9	14.4	26.9	0.88	1.59
WAIS-R Vocabulary	0.20	0.08	1.69	12.6	16.5	30.8	1.31	2.44**
WAIS-R Comprehension	0.10	0.13	2.12	13.4	15.5	30.8	1.15	2.30*
WAIS-R Block Design	1.44	1.70	1.92	24.8	25.8	26.9	1.03	1.06
WAIS-R Digit Symbol	1.35	1.88	3.27	20.1	26.8	42.3	1.33	2.11**
MMSE	0.01	1.21	3.38	15.0	29.9	38.5	2.00**	2.57**
Verbal Fluency	0.14	0.70	2.85	14.2	17.5	42.3	1.24	2.99***
Word List Recall	0.19	1.39	3.27	15.0	30.9	42.3	2.07***	2.83***
Praxis Delayed Total	1.06	1.86	2.92	23.6	29.9	38.5	1.27	1.63
Trails A	1.37	1.78	2.46	21.3	29.9	34.6	1.41	1.63
Trails B	1.31	1.82	2.92	21.3	27.8	38.5	1.31	1.81*

* $p < .05$; ** $p < .01$; *** $p < .001$

When extended into the psychometric abilities factor structure, the neuropsychology measures generally spread over two or more of the psychometric abilities. However, examination of the primary (i.e., largest) loading for each test revealed a somewhat more simple explanation of the factors. Six tests had their highest loading on the Verbal Memory factor, and five of these are primarily identified as memory tests: Word List Recall, WMS-R Immediate and Delayed Recall Fuld Retrieval and Rapid Verbal Retrieval. The sixth test, the MMSE, also has a strong verbal memory component. A different set of six tests loaded most strongly on the Spatial Orientation factor, and five of these six have a significant spatial ability component: Constructional Praxis Delayed, WAIS-R Digit Symbol and Block Design, and Trails A and B. The sixth test on this factor, the Boston Naming Test, while often found to load on spatial orientation, is primarily a verbal ability measure. Six tests had their highest loading on the Verbal Comprehension factor, although the loading for the WMS-R Immediate Recall on this factor was equal to that for the Verbal Memory factor. Three of these tests were primarily verbal ability tests: Verbal Fluency, WAIS-R Vocabulary and WAIS-R Comprehension. Although the WAIS-R Digit Span and WMS-R Immediate had high loadings on the Verbal Comprehension factor, they also had equal or nearly equal loadings on other factors–the Inductive Reasoning and Verbal Memory, respectively. These double loadings indicate that multiple mental abilities are implicated in performance on most neuropsychological assessment instruments. Most measures had a secondary loading on the Spatial Orientation factor, except for the MMSE, the WAIS-R Digit Span scale, and the Wechsler Memory Immediate Recall. Several measures also had secondary and/or tertiary loadings on the Inductive Reasoning and Numeric Ability factors.

These findings suggest that, at least for the main components of the neuropsychological battery, we may be able to predict substantial proportions of variance from our psychometric ability battery. It then becomes possible to use our longitudinal psychometric measures to obtain estimates of what the status of our study participants on neuropsychological measures might have been at earlier points in time, had we been able to administer such measures directly. In contrast to the highly skewed distribution of actually observed neuropsychology tests found in a normal population, the estimated scores exhibit a psychometrically much better behaved distribution. Hence, we would suggest that it is such estimated scores that should be used to determine early signs of impairment or estimates of risk in normal populations.

We first evaluated this approach by obtaining concurrent estimates of the neuropsychology tests from the primary mental ability measures. This analysis provides neuropsychology tests from the primary mental ability measures. This analysis provides estimates that are very close to the proportion of reliable variance in these tests. We then proceeded to estimate

measures that might have been obtained seven and fourteen years earlier respectively. Findings indicate that for our community-dwelling sample, age-related declines occurred over 14 years in all age groups (except for the Mattis scale), and for a few neuropsychological measures over 7 years in the old-old and very-old age groups.

A major criterion for the utility of the analyses presented here is, of course, whether the backward estimation of neuropsychological measures can contribute to the detection of potential risk of dementia at an earlier point in time when the direct identification by a neuropsychological battery would not be practical because of expected ceiling effects. We therefore validated our approach applying the criterion of a well-established procedure of cognitive impairment consensus ratings used by neuropsychologists.

Our results suggest first that significant individual change on primary mental ability test performance over the 7 years preceding the neuropsychological evaluation has predictive value for identifying individuals who will be rated by neuropsychologists to be cognitively impaired. More importantly, while there is some predictability directly from the psychometric test battery, there is better prediction if we derive estimated neuropsychology test scores. Furthermore, we can also successfully predict current diagnostic status from change in the estimated neuropsychology measures from 14 years to 7 years prior to the actual administration of the neuropsychology battery.

As indicated above, a major advantage of the approach here taken is that, in contrast to the actual neuropsychology tests, the estimated scores have no ceiling since they are scaled from the midpoint of the total normal population. Removing the ceiling limitation for the estimated neuropsychology tests, of course, makes it possible for the estimated neuropsychology scores to show greater predictive efficacy than the direct measures of change in the primary mental abilities. While the psychometric factor scores provide good overall status measures on level of cognitive functioning, they were not designed to be specifically relevant to the detection of neuropathology. That is, they were constructed to reflect basic mental abilities, while the neuropsychology measures were constructed to detect neuropathology. The predicted neuropsychology scores, on the other hand, not only have these desired attributes but also have the psychometric characteristics required for the assessment of risk of dementia in normal population, as well as for studying longitudinal change in these characteristics.

Efforts to develop programs for the prevention or arrest of dementia at early stages will depend heavily on the early identification of those at risk *before* clinical symptoms begin to appear. In this article we have presented a novel approach that takes advantage of existing longitudinal data to identify individuals at risk by post-dicting performance on neuropsychological tests seven and fourteen years prior to neuropsychological assessment. Given the

increasing availability of longitudinal data bases that could be related to eventual occurrence of dementia this approach is thought to have considerable promise. Finally, it may be suggested that the approach described here is useful also across many subject areas when new measurement instruments are added to a longitudinal study, and when it may be important to explore how study participants might have performed on the new measures had they been availablein the past.

AUTHOR NOTE

We wish to thank the staff and members of the Group Health Cooperative of Puget Sound for their long-term support and participation in the Seattle Longitudinal Study. We also wish to acknowledge the neuropsychology testers, Charles Fick, Michael Matushka, Holly Overman and Sara Rosen who were trained under the supervision of Linda Teri. Our colleagues and co-authors Alfred Kaszniak and Linda Teri contributed the expert ratings of cognitive impairment. Further information on the Seattle Longitudinal Study can be obtained at http://geron.psu.edu/sls. Inquiries regarding this article should be addressed to K. Warner Schaie, Gerontology Center, Pennsylvania State University, 135 E. Nittany Ave., Suite 405, State College, PA 16801; e-mail: kws@psu.edu.

REFERENCES

Arbuckle, J. L., & Wothke, W. (1999). *Amos 4.0 User's Guide.* Chicago, IL: Small Waters Corporation.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, *221*, 82–90.

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensorand cognitive functions across the adult life span: A new window to the study of cognitive aging. *Psychology and Aging*, *12*, 12–21.

Blieszner, R., Willis, S. L., & Baltes, P. B. (1981). Training research in aging on the fluid ability of inductive reasoning. *Journal of Applied Developmental Psychology*, *2*, 247–265.

Bondi, M., Salmon, D.., Monsch, A., Galasko, D., Butters, N., Klauber, M., Tahl, L., & Saitoh, T. (1995). Episodic memory changes are associated with the ApoE-4 allele in non-demented older adults. *Neurology*, *45*, 2203–2206.

Cahn, D. A., Salmon, D. P., Butters, N., Wiederholt, W. C., Corey-Bloom, J., Edelstein, S. L., & Barrett-Connor, E. (1995). Detection of dementia of the Alzheimer type in a population-based sample: Neuro-psychological test performance. *Journal of the International Neuropsychological Society*, *1*, 252–260.

Crum, R. M., Anthony, J. C., Bassett, S. S., & Folstein, M. F. (1993). Population-based norms for the Mini-Mental State Examination by age and education level. *Journal of the American Medical Association*, *269*, 2386–2391.

Dwyer, P. S. (1937). The determination of the factor loadings of a given test from the known factor loadings of other tests. *Psychometrika*, *2*, 173–178.

Ekstrom, R. B., French, J. W., Harman, H., & Derman, D. (1976). *Kit of factor-referenced cognitive tests* (rev. ed.). Princeton, NJ: Educational Testing Service.

Epstein, N. B., Baldwin, L. M., & Bishop, D. S. (1983). The McMaster family assessment device. *Journal of Marital and Family Therapy*, *9*, 171–180.

Epstein, N. B., Bishop, D., Ryan, C., Miller, I., & Keitner, G. (1993). The McMaster model: View of health family functioning. In F. Walsh (Ed.), *Normal family processes.* New York: Guilford.

Folstein, M. F., Folstein, S.E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.

Fuld, P. A. (1977). *Fuld Object-Memory Evaluation.* Woodsdale, IL: Stoelting Co.

Garrett, H. E. (1966). *Statistics in psychology and education* (6th ed.). New York: McKay.

Gilewski, MJ., & Zelinski, E. M. (1988). The Memory Functioning Questionnaire (MFQ). *Psychopharmacology Bulletin*, *24*, 665–670.

Golden, C.J., & Chronopoulos, A. (1998). Dementia. In M. Hersen, & V. B. Van Hasselt (Eds.), *Handbook of clinical geropsychology* (pp. 113–146). New York: Plenum.

Green, R. C., Woodard, J. L., & Green, J. (1995). Validity of the Mattis Dementia Rating Scale in detection of cognitive impairment in the elderly. *Journal of Neuropsychiatry*, *7*, 357–360.

Heaton, R. K., Grant, I., & Matthews, C. G. (1986). Differences in neuropsychological test performance associated with age, education, and sex. In I. Grant, & K. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (pp. 10–120). New York: Oxford Press.

Horn, J. L. (1982). The theory of fluid and crystallized intelligence in relation to concepts cognitive psychology and aging in adulthood. In F. I. M. Craik, & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 237–278). New York: Plenum.

Isaacs, B., & Kennis, A. T. (1973). The set test as aid to the detection of dementia in old people. *British Journal of Psychiatry*, *123*, 467–470.

Kaplan, E., Goodglass, H., & Weintraub, S. (1984). *The Boston Naming Test.* Philadelphia, PA: Lee & Febinger.

Kaufman, A. S., Kaufman, J. L., McLean, J. E., & Reynolds, C. R. (1991). Is the pattern of intellectual growth and decline across the adult life span different for men and women? *Journal of Clinical Psychology*, *47*, 801–820.

Kiyak, H. A., Teri, L., & Borson, S. (1994). Physical and functional health assessment in normal aging and in Alzheimer's disease: Self-reports vs. family reports. *Gerontologist*, *34*, 324–330.

Lafleche, G., & Albert, M. (1995). Executive function deficit in mild Alzheimer's disease. *Neuropsychology*, *9*, 313–320.

LaRue, A. (1992). *Aging and neuropsychological assessment: Critical issues in neuropsychology.* New York: Plenum.

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist*, *9*, 179–185.

Lemsky, C. M., Smith, G., Malec, J. R., & Ivnik, R. J. (1996). Identifying risk for functional impairment using cognitive measures: An application of CART modeling. *Neuropsychology*, *10*, 42–52.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Matarazzo, J. D. (1972). *Wechsler's measurement and appraisa l of adult intelligence.* Baltimore: Williams & Wilkins.

Mattis, S. (1988). *Dementia rating scale.* Odessa, FL: Psychological Assessment Resources.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525–543.

Mohs, R. C., Kim, Y., & Johns, C. A. et al. (1986). Assessing change in Alzheimer's disease: Memory and language tests. In L. W. Poon (Ed.), *Handbook for clinical memory assessment of older adults* (pp. 149–155). Washington, DC: American Psychological Association.

Monsch, A., Bondi, M., Salmon, D., Butters, N., Tahl, L., Hansen, L., Wiederholt, W., Cahn, D., & Klauber, B. (1995). Clinical validity of the Mattis Dementia Rating Scale in detecting dementia of the Alzheimer's type. *Archives of Neurology*, *52*, 899–904.

Morris, J. C., Edland, S., Clark, C., Galasko, C., Koss, E., Mohs, R., Van Belle, G., Fillenbaum, G., & Heyman, A. (1993). The consortium to establish a registry for Alzheimer's disease (CERAD). Pare IV. Clinical and neuro-psychological assessment of Alzheimer's disease. *Neurology*, *39*, 1159–165.

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., Mellits, Ed.D., & Clark, C. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD). Part I. Rates of cognitive change in the longitudinal assessment of probable Alzheimer's disease. *Neurology*, *43*, 2457–2465.

Radloff, L. S. (1977). A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *3*, 385–401.

Radloff, L. S., & Teri, L. (1986). Use of the center for epidemiological studies-depression scale with older adults. *Clinical Gerontologist*, *5*, 119–137.

Reitan, R. M., & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation.* Tucson, AZ: Neuropsychology Press.

Rosen, W. G., Mohs, R. C., & Davis, K. L. (1984). A new rating scale for Alzheimer's diseases. *American Journal of Psychiatry*, *14*, 1356–1364.

Ryan, J. J., Paolo, A. M., & Brugardt, T. M. (1990). Test-retest stability of the WAIS-R in normal subjects 75 years and older. *Journal of Clinical Experimental Neuropsychology*, *12*, 58.

Ryan, J. J., Paolo, A. M., & Brugardt, T. M. (1992). WAIS-R test-retest stability in normal subjects 75 years and older. *Clinical Neuropsychologist*, *6*, 3–8.

Schaie, K. W. (1985). *Manual for the Schaie-Thurstone Adul t Mental Abilities Test (STAMAT).* Palo Alto, CA: Consulting Psychologists Press.

Schaie, K. W. (1989). Perceptual speed in adulthood: Cross-sectional and longitudinal studies. *Psychology and Aging*, *4*, 443–453.

Schaie, K. W. (1996). *Intellectual Development in Adulthood: The Seattle Longitudinal Study.* New York: Cambridge University Press.

Schaie, K. W. (2000). Longitudinal and related methodological issues in the Swedish Twin Registry. In B. Smedby, I. Lundberg, & T. I. A. Sørensen (Eds.), *Scientific evaluation of the Swedish Twin Registry* (pp. 62–74). Stockholm, Sweden: Swedish Council for Planning and Coordination of Research.

Schaie, K. W. (2004). *Developmental influences on adult intelligence: The Seattle Longitudinal Study.* New York: Oxford University Press.

Schaie, K. W., Dutta, R., & Willis, S. L. (1991). The relationship between rigidity-flexibility and cognitive abilities in adulthood. *Psychology and Aging*, *6*, 371–383.

Schaie, K. W., Maitland, S. B., Willis, S. L., & Intrieri, R. L. (1998). Longitudinal invariance of adult psychometric ability factor structures across seven years. *Psychology and Aging*, *13*, 8–20.

Smith, G. E., Ivnik, R. J., Malec, J. F., Kokman, E., Tangelos, E., & Petersen, R. C. (1994). Psychometric properties of the Mattis Dementia Rating Scale. *Psychological assessment*, *2*, 123–131.

Spreen, O., & Strauss, E. (1991). *A compendium of neuropsychological l tests: Administration, norms, and commentary.* New York: Oxford University Press.

Storandt, M., Botwinick, J., Danziger, W., Berg, L., & Hughes, C. (1984). Psychometric differentiation of mild senile dementia of the Alzheimer's type. *Archives of Neurology*, *41*, 497–499.

Thurstone, L. L. (1938). *The primary mental abilities.* Chicago: University of Chicago Press.

Thurstone, L. L., & Thurstone, T. G. (1949). *Examiner Manual for the SRA Primary Mental Abilities Test (Form 10–14).* Chicago: Science Research Associates.

Tombaugh, T. N., & McIntyre, N. J. (1992). The MMSE: A comprehensive review. *Journal of the American Geriatric Society*, *40*, 922–935.

Tucker, L. R. (1971). Relations of factor score estimates to their use. *Psychometrika*, *36*, 427–436.

Uhlmann, R. F., & Larson, E. B. (1991). Effect of education on the Mini-Mental Status Examination as a screening test for dementia. *Journal of the American Geriatric Society*, *39*, 876–880.

Vangel, S. J., & Lichtenberg, P. Q. (1995). Mattis Dementia Rating Scale: Clinical utility and relationship with demographic variables. *Clinical Neuropsychologist*, *9*, 209–213.

Van Gorp, W. G., Satz, P., Kiersch, M. E., & Henry, R. (1986). Normative data on the BNT for a group of normal older adults. *Journal of Clinical Experimental Neuropsychology*, *8*, 702–705.

Vinters, H. V. (2001). Aging and the human nervous system. In J. E. Birren, & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 135–160). San Diego, CA: Academic Press.

Vitaliano, P., Russo, J., Breen, A., Vitello, M., & Prinz, P. (1986). Functional decline in the early stages of Alzheimer's disease. *Psychology and Aging*, *1*, 41–46.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-revised.* Cleveland, OH: Psychological Corporation.

Welsh, K. A., Butters, N., Mohs, R. C., Beekly, D., Edland, S., Fillenbaum, G., & Heyman, A. (1994). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. *Neurology*, *44*, 609–614.

Wilcox, G. K. (2003). Alzheimer's disease. In R. C. Tallis, & H. M. Fillit (Eds.), *Geriatric Medicine and gerontology* (6th ed., pp. 797–806). London, UK: Churchill Livingstone.

Woodruff, -Pak D. S., & Papka, M. (1999). Theories of neuropsychology and aging. In V. L. Bengtson, & K. W. Schaie (Eds.), *Handbook of theories of aging* (pp. 113–132). New York: Springer Publishing Co.

Zelinski, E. M., Gilewski, M. J., & Schaie, K. W. (1993). Three-year longitudinal memory assessment in older adults: Little change in performance. *Psychology and Aging*, *8*, 176.