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

 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.

1

Introduction

OVERVIEW

The Neuropsychological Assessment Battery (NAB) is a comprehensive, modular battery of neuropsychological tests developed for the assessment of a wide array of cognitive skills and functions in adults aged 18 years to 97 years, with known or suspected disorders of the central nervous system. The NAB was created over a 7-year period, funded, in part, through grants from the United States National Institute of Mental Health (1 R43 MH58501-01, 2 R44 MH58501-02). Decisions pertaining to the content and format of the NAB were guided by the results of a national survey of neuropsychological assessment needs (Stern & White, 2000; see chapter 2) and by the feedback and guidance from members of the NAB Advisory Council and other consultants (see Appendix A). As described later in this chapter, 10 innovative features were incorporated into the design of the NAB. These features are presented in Table 1.1. The NAB includes psychometrically sound quantitative summary indexes, along with numerous scores that address the qualitative or "process" features of performance. The NAB was nationally standardized on 1,448 adults ranging in age from 18 to 97 years. Although normed as a complete battery, the NAB also allows for flexibility in test administration with regard to both time and areas of focus (see chapter 1 of the *NAB Administration, Scoring, and Interpretation Manual*, Stern & White, 2003).

Table 1.1
Innovative Features of the NAB

Feature
Screening for both impaired and normal performance ("dual-screening capability")
Comprehensive coverage of functional domains
Combined strengths of flexible and fixed battery approaches to assessment
Avoidance of floor and ceiling effects
Reduced administration time
Entire battery normed on a single standardization group ("coordinated norming")
Demographically corrected norms based on age, education level, and sex
Provision of an equivalent/parallel form
Increased user-friendliness for both examiner and examinee
Focus on ecological validity

HISTORICAL PERSPECTIVE

Neuropsychological assessment is integral to multidisciplinary health care (American Academy of Neurology, 1996; Levin, 1994), and neuropsychologists contribute valuable information regarding clinical diagnosis, patient care and disposition planning, rehabilitation and treatment evaluation, and basic and applied research (Lezak, 1995). The importance of neuropsychology is manifest in numerous ways, including (a) the increasing number of scientific journals devoted to the field; (b) the founding of advanced neuropsychological credentialing bodies (e.g., American Board of Clinical Neuropsychology, ABCN; American Board of Professional Neuropsychology, ABPN); and (c) the increased number of doctoral programs, internships, and postdoctoral fellowship training sites providing specialized training in neuropsychology (Meier, 1992). In addition, both the American Psychological Association (APA) and the Canadian Psychological Association (CPA) formally recognize clinical neuropsychology as a distinct specialty area within applied psychology. Furthermore, neuropsychological evaluations have become an important component of criminal and civil legal proceedings (Faust, Ziskin, & Hiers, 1991).

In the past, neuropsychological test batteries were used mainly because they were superior diagnostic alternatives to the pneumoencephalograph and first generation brain scans (Long, 1996). Early batteries, such as the Halstead-Reitan Neuropsychological Battery (HRNB; Reitan & Wolfson, 1993) and the Luria-Nebraska Neuropsychological Battery (LNNB; Golden, Purisch, & Hammeke, 1985), provided

estimates about the presence, laterality, localization, and extent of cerebral damage. Administration was often lengthy and complicated, but in the pre-managed care era of third-party reimbursement, the healthcare marketplace supported neuropsychologists who offered these services.

More recent developments now pose significant challenges to the field of clinical neuropsychology. For example, newer generation structural and functional neuroimaging techniques (e.g., computerized tomography [CT], magnetic resonance imaging [MRI], position emission tomography [PET], functional magnetic resonance imaging [fMRI]) have significantly reduced neuropsychology's diagnostic/localization role and replaced it with new evaluation goals. Clinicians are now asked to address issues of cognitive retraining, compensatory strategies, adaptive skills, treatment options, rehabilitation potential, and optimal living arrangements (Heinrichs, 1990). These new goals require new assessment methods (Long, 1996). Furthermore, changes in the economics of health care delivery and reimbursement have shifted the focus of services toward brief and essential medical tests and interventions. In many situations, these economic changes require neuropsychologists to limit the time spent in evaluating their patients. Many clinicians have responded by cobbling together their own individualized batteries of separate tests that have variable psychometric validity (Lezak, 1995; Spreen & Strauss, 1998); the psychometric functioning of these tests as an integrated unit is simply unknown (Russell, 1994) and, in many cases, are not recognized as meeting the legal criteria of an acceptable battery of instruments in forensic settings (Ziskin, 1995). Such individualized batteries usually do not have screening versions, do not have parallel alternate forms, contain tests that are normed and validated on different samples, are relatively expensive, and lack "ecological" validity.

Theoretical Models of Neuropsychological Assessment

From the historical roots of neuropsychology, two approaches to neuropsychological assessment have emerged. The first, labeled the cognitive/metric (Russell, 1994) or fixed battery (Larrabee, 2000) approach, is exemplified by the HRNB and the LNNB. The other, labeled the flexible battery (Bauer, 1994) approach, is exemplified by the work of Lezak (1995) and Walsh (1987). These two orientations differ from each other with respect to the nature and timing of test-selection decisions and the relative reliance on psychometric versus neurologic concepts in conceptualizing the process and goals of assessment (Bauer).

Each approach has its own combination of strengths and limitations, and practitioners are constantly trying to strike a balance between administration time, data yield, and diagnostic requirements when applying the respective assessment model. A fixed battery provides a constant background of tests (Russell, 1994), highly standardized administration and scoring procedures, quantitative data (i.e., scores on an interval-level scale), and the ability to accumulate an impressive body of validation research (e.g., Boll, 1981; Filskov & Goldstein, 1974). However, fixed batteries require unnecessarily long administration time, in that all patients receive the same exhaustive set of instruments, regardless of the specific diagnostic question, goals of the evaluation, or individual strengths and weaknesses. Conversely, although they are more patient-focused, flexible batteries do not provide the advantages of a constant background of tests and place much heavier demands on subjective clinical judgment and knowledge of tests and brain–behavior relationships than the fixed battery approach (Lezak, 1995). Clinicians using flexible batteries often choose to limit the functional domains of behavior that are assessed as a means of reducing administration time. These decisions create a significant risk of obscuring other viable hypotheses for observed impairments (Bauer, 1994).

Implementation of Neuropsychological Batteries

In an attempt to deal with the design and psychometric flaws of existing fixed batteries of tests and also to maximize the data-yield-to-administration-time ratio, Bauer (1994) has advocated the use of multiple fixed batteries to meet the particular demands of homogeneous patient groups, differential diagnoses, and referral questions and sources. Bauer identifies three types of multiple fixed batteries: (a) a general screening battery, (b) a population-specific battery, and (c) a domain-specific battery. *General screening batteries* contain a small number of highly sensitive items that tap complex, multifactorial neuropsychological functions and should measure a breadth, but not depth, of cognitive domains. Benton (1992) suggested that this type of screening battery should require no more than 60 minutes to administer and score. The goal of screening is to determine the presence versus absence of abnormal/impaired performance and the need for additional follow-up testing. *Population-specific batteries* contain tests that measure functions known or suspected to be relevant for a particular diagnosis or disease, such as epilepsy (Dodrill, 1978), multiple sclerosis (Peyser, Rao, LaRocca, & Kaplan, 1990), and Alzheimer's dementia (Welsh, Butters, Hughes, Mohs, & Heyman, 1992). These batteries may provide specialists in these areas important clinical and research information pertaining to the specific disorder but do not allow for broader evaluation goals or comparison of findings across different diagnostic groups.

Domain-specific batteries contain tasks designed to provide for a detailed assessment of a particular functional domain. Examples of domain-specific batteries include the Wechsler Memory Scale, Third Edition (WMS-III; Wechsler, 1997b), the Boston Diagnostic Aphasia Examination, Third Edition (BDAE; Goodglass, Kaplan, & Baressi, 2000), and the Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2001). These batteries allow for more in-depth assessment of a specific area of functioning although they frequently require more time to administer than is often available, especially when several domain-specific batteries are combined for a single evaluation.

Tarter and Edwards (1986) proposed a three-stage, "step" model of neuropsychological assessment to guide the implementation of multiple fixed batteries in clinical practice. The three steps include screening, additional testing with specific batteries, and "idiographic testing." The first step involves administration of highly sensitive screening tasks to all patients, regardless of their presumed diagnoses and capabilities. Patients who pass difficult screening items would not need further testing. Patients who fail complex screening items would trigger Step 2, the decision to order additional testing. Depending upon the items failed in Step 1, additional testing might involve (a) a single, domain-specific battery, (b) several domain-specific batteries, or (c) a population-specific battery. Based on the results of the second tier of testing, the decision is made either to terminate testing or to proceed to Step 3, idiographic testing, which is individually tailored to the specific test findings and to the specific characteristics of the referral question.

RATIONALE FOR THE NAB

The field of neuropsychology has lacked an integrated battery of assessment instruments capable of providing sophisticated test data while requiring only a relatively brief administration time (Benton, 1992). To meet current needs, such a battery should (a) have good psychometric characteristics, (b) include extensive normative and validation data, (c) provide clinical information that satisfies a broad range of modern referral sources and questions, and (d) facilitate systematic research (Parsons, 1993). The goal underlying the development of the NAB was to address these needs with a new and innovative neuropsychological test battery that provides a comprehensive evaluation of neuropsychological functions in less than 4 hours. The NAB incorporates the conceptual framework of Bauer (1994) and Tarter and Edwards (1986) by offering a separate Screening Module to indicate the need to administer the main NAB modules. For those areas of idiographic testing not included in the NAB (e.g., motor functioning, mood/personality), the examiner can expand upon the NAB assessment with his/her favored instruments. Moreover, the individual examiner may choose to forego the Screening Module and administer any or all of the five main modules to a patient, on the basis of specific clinical needs. In addition, the flexibility inherent in the battery also allows for selection of individual tests from each module, rather than requiring the administration of an entire main NAB module, when this type of non-battery-focused assessment is clinically warranted. Table 1.2 presents the six NAB modules.

Table 1.2
NAB Modules

Module	Acronym
Screening	SCR
Attention	ATT
Language	LAN
Memory	MEM
Spatial	SPT
Executive Functions	EXE

Survey of Neuropsychological Assessment Practices

In order to ascertain the needs of the potential users of a new neuropsychological test battery, the publisher conducted a comprehensive national survey of neuropsychological assessment practices (Stern & White, 2000), which was the first of its kind in the field of neuropsychology. The results served as a basis for the development of the NAB with regard to issues such as assessment areas of functional domains, total length, and other salient content and format characteristics of the battery. The survey population was identified through membership lists of the three professional organizations in the field: the International Neuropsychological Society (INS), the National Academy of Neuropsychology (NAN), and Division 40 (Clinical Neuropsychology) of the American Psychological Association (APA). Membership lists were merged to form a master list and purged of duplicate entries and individuals residing outside the United States. In June 1997, a total of 7,388 surveys were mailed with postage-paid return envelopes. Respondents were asked to provide anonymous information about their demographic and educational background, their current practice, and the effects of managed care and other reimbursement issues on their practice. Respondents also gave opinions on the need for a new, briefer-yet-comprehensive neuropsychological assessment battery, desirable properties of this new instrument, and functional neuropsychological domains that would be important to include in such an instrument.

A total of 1,011 (15%) surveys were returned; however, respondents who evaluate only children (n = 112; 11% of returned surveys) or who did not indicate a client base (n = 11; 1% of returned surveys) were excluded from the data analyses. The vast majority of the remaining respondents (93%) hold PhD or PsyD degrees and characterize themselves as either neuropsychologists (70%) or clinical psychologists (21%). Only a small proportion of the respondents were diplomates of either ABCN (9%) or ABPN (5%). However, these figures represent a significant proportion of the membership of these credentialing bodies (e.g., 35% of the 286 ABCN diplomates at that time returned the survey). Professional practice settings included independent private practices (39%) and rehabilitation facilities, hospitals, and academic medical centers (44% combined). Fifty-two percent of respondents reported devoting more than 50% of their professional time to clinical neuropsychological services, and 39% reported not devoting any time to research. Forty-eight percent utilized neuropsychological tests for clinical evaluations *only*, whereas 50% used tests for *both* research purposes and clinical evaluations. When asked about their approach to neuropsychological test selection, 68% of respondents reported using a *flexible battery* approach (i.e., "variable but routine grouping of tests"), 17% reported using a *flexible* approach (i.e., "based on needs of the individual patient"), and 13% reported using a *standardized battery* (i.e., "routine grouping of tests"; HRNB, LNNB).

Respondents rated whether managed care and/or other changes in funding have resulted in the need to conduct shorter neuropsychological evaluations. Only a small minority (15%) endorsed the statement as false, not at all true, whereas 85% said the statement was true to some extent (34% slightly true, 28% mainly true, and 23% very true). An important finding of the survey was the significant discrepancy between the amount of time respondents thought was ideally needed for a comprehensive neuropsychological evaluation given current instrumentation (*Mode* = 5 to 6 hours; 25% stated 4 hours or less) and the amount of time they thought was required to conduct a realistic and reimbursable neuropsychological evaluation in today's health care climate (*Mode* = 3 to 4 hours; 49% stated 4 hours or less). Eighty-nine percent of respondents stated that, at the time of the survey, there was no commercially available instrument that provided a comprehensive evaluation within the current time/funding constraints.

Survey respondents were provided with 36 specific characteristics of a "brief, comprehensive clinical neuropsychological instrument" and asked to rate their importance on a 4-point scale (from 1 = *Not at all important* to 4 = *Very*

Table 1.3
Survey Ratings of the Importance of Various Characteristics for a New Comprehensive Neuropsychological Test Battery

Test characteristic	Not at all important	Slightly important	Moderately important	Very important
Demographically corrected norms	1.3	3.5	18.0	77.2
Detects changes in cognitive functioning across time	0.8	3.2	32.8	63.1
May be used in entirety or as separate modules	2.0	8.4	29.2	60.4
Appropriate for high functioning patients (avoids ceiling effects)	0.9	9.2	33.2	56.7
Provides several impairment indexes for specific functions	4.1	11.6	37.5	46.9
Ecologically valid	7.2	13.7	34.4	44.7
Submissible as evidence in forensic cases	7.2	17.4	30.9	44.4
Ethnic and minority group norms	4.0	16.1	40.0	39.9
Appropriate for severely impaired patients (avoids floor effects)	6.2	21.4	37.1	35.4
Has screening items that quickly rule out intact patients	7.0	18.1	40.0	34.8
Provides qualitative or process scoring and interpretation	7.4	26.4	38.3	27.9
Computerized norming and scoring	16.9	25.4	30.4	27.3
Repeatable and equivalent forms	3.6	27.7	41.8	26.9
May be administered by a psychometrician or research assistant	25.8	33.6	25.5	13.1
Computerized interpretation reports	40.6	30.8	18.7	9.9
Computerized administration	37.5	35.0	17.8	9.7

Note. N = 888. Source: PAR Survey of Neuropsychological Assessment Needs (Stern & White, 2000). Test characteristics are listed in descending order of percentage rated as very important.

important). Table 1.3 presents the results for 16 of the characteristics that were rated as most important. It is noteworthy that "use of computer for administering at least some items or tasks" and "provides a single summary impairment index" were rated poorly (i.e., only 27% and 32% of the respondents, respectively, rated these aspects as either moderately or very important).

Innovative Features of the NAB

The results of the publisher's survey of neuropsychological assessment needs led to the inclusion of 10 innovative features in the NAB (Stern & White, 2000). Each of these features is discussed in the following sections, along with relevant survey results (see Table 1.1).

"Dual-Screening Capability"

Screening capability was rated as moderate-to-very important by 75% of the survey respondents. In practice, neuropsychological screening is typically geared toward identifying patients who show no signs of brain dysfunction and no need for extensive follow-up testing. This approach has been formally incorporated into two popular assessment instruments, the Dementia Rating Scale–Second Edition (DRS-2; Jurica, Leitten, & Mattis, 2001) and Cognistat (also known as the Neurobehavioral Cognitive Status Examination [NCSE]; Kiernan, Mueller, & Langston, 1987). The NAB incorporates this screening algorithm for each functional domain assessed but *also extends this capability to screen out patients who are too impaired to benefit from additional testing*. If the referral question requires qualification and description of the patient's functioning, the user can always disregard the screening recommendation and administer the entire battery or selected main module(s).

Comprehensive Coverage of Functional Domains

Reviews of the neuropsychological literature (e.g., Lezak, 1995; Mapou & Spector, 1995; Spreen & Strauss, 1998) have identified seven major functional domains: (a) language and verbal communicative functions; (b) spatial/perceptual skills; (c) sensorimotor functions; (d) attention and related information processing tasks (including working memory); (e) learning and memory; (f) executive functions and problem-solving abilities; and (g) personality, emotional, and adaptive functions. This conceptual framework has been confirmed with factor analytic studies of various neuropsychological batteries (Larrabee & Curtiss, 1992; Leonberger, Nicks, Larrabee, & Goldfader, 1992). Many neuropsychologists also add to their evaluations measures of overall intellectual functioning and, especially in forensic settings, measures of malingering/symptom validity. The NAB was developed with the overriding goal of providing a common set of core tests that serve as a reasonably comprehensive standard reference base suitable for most routine clinical applications. Thus, the NAB is specifically neither a "screening battery" nor an exhaustive test battery that measures every conceivable neuropsychological skill and related functions. The survey of neuropsychologists led to decisions pertaining to the final content composition of the NAB and lends strong support for organizing the NAB into six modules: Screening, Attention, Language, Memory, Spatial, and Executive Functions.

Combined Strengths of Flexible and Fixed Battery Approaches to Assessment

As described previously, there are both strengths and limitations to each of the existing major approaches to neuropsychological assessment (i.e., flexible battery and fixed battery). In the development of the NAB, the strengths of each of these approaches were included. Therefore, the NAB provides the following features: (a) a constant background of tests, with a focused, patient-centered examination and shorter administration times afforded by the dual-screening approach; (b) standardized administration and scoring procedures across all tests; (c) quantitative summary indexes along with numerous measures of pertinent qualitative aspects of performance; and (d) minimal reliance on clinical decision making in test selection. This overall approach also allows for the accumulation of extensive validation research.

Avoidance of Floor and Ceiling Effects

Approximately 90% of survey respondents indicated that it would be moderately or very important for a new comprehensive test battery to be appropriate for high functioning examinees and should, therefore, avoid ceiling effects. Although not as highly rated (i.e., 73% giving a rating of moderately or very important), survey respondents indicated that a new battery should also be appropriate for severely impaired patients and should, therefore, avoid floor effects. A guiding principle in the development of the NAB was the avoidance of both ceiling and floor effects, when appropriate. For most tests in the NAB, a continuum of difficulty levels was included to provide a relatively normal distribution in test performance. Difficulty ratings were provided by the Advisory Council members and used in the initial creation and selection of individual test items. In addition, difficulty analyses were conducted on data collected at both pilot testing and standardization to ensure the adequacy of distributions.

Reduced Administration Time

The NAB provides a reasonably comprehensive evaluation in a much briefer time period than is currently available. Approximately 71% of the survey respondents indicated that

a realistic and reimbursable neuropsychological evaluation can be completed within 3 to 4 or 4 to 5 hours (excluding record review, interviewing, and report writing). The entire NAB requires approximately 3 hours for the five main modules and less than 4 hours for all six modules (Screening Module and five main modules). Table 1.4 presents the approximate administration time for each module and the total battery. In most situations, clinicians still have time to administer intelligence and personality tests, as well as to pursue idiographic testing (e.g., motor skills, effort testing) when clinically warranted.

Table 1.4
Approximate Administration Time for the NAB Modules

Module	Administration time
Screening	45 minutes
Attention	45 minutes
Language	35 minutes
Memory	45 minutes
Spatial	25 minutes
Executive Functions	30 minutes
Full NAB (5 main modules)	180 minutes (3 hours)
Screening Module and Full NAB	220 minutes (3 hours, 40 minutes)

Coordinated Norming

Whereas much is known about the psychometric properties of individual neuropsychological tests (Franzen, 1989; Lezak, 1995; Mitrushina, Boone, & D'Elia, 1998; Spreen & Strauss, 1998), very little effort has been devoted to the examination of how individual instruments function within a battery (Russell, 1994). Given the fact that 85% of the survey respondents reported using a *customized battery*, the lack of psychometric data on customized batteries represents a very large gap in the neuropsychological knowledge base and may lead to critical limitations in the overall validity of clinical decisions based on neuropsychological test data (Faust et al., 1991). In fact, this lack of coordinated norming of customized batteries may render forensic examination results based on these tests inadmissible as evidence in court according to the Daubert ruling (Ziskin, 1995). It is important to note that 75% of the survey respondents rated "submissible as evidence in forensic cases" as either moderately important or very important. In addition, because of the potential for specific neuropsychological test performance to be associated with IQ level (Tremont, Hoffman, Scott, & Adams, 1998), it is important to understand and quantify the relationship between a battery of neuropsychological tests and a measure of overall IQ. The NAB fills these gaps by providing coordinated norms for all of the NAB modules, along with a recently published measure of intelligence, the Reynolds Intellectual Screening Test (RIST; Reynolds & Kamphaus, 2003). The RIST is an excellent measure of general intelligence (*g*) and correlates highly with the Full Scale IQ of the Wechsler Adult Intelligence Scale, Third Edition (WAIS-III; Wechsler, 1997a). These coordinated norms allow for within- and between-patient score comparisons across the NAB and between these specific measures and estimated IQ level. Moreover, *the examiner can use a single set of normative tables* (including appropriate age, sex, and education corrections) for the entire NAB, rather than dealing with the commonly used mixture of test-specific norms compiled (often uniquely by each examiner) in each examiner's "norms book."

Demographically Corrected Norms

The need to interpret neuropsychological tests in the context of an individual's age, educational attainment, and sex has been well established in the field (Heaton, Grant, & Matthews, 1991). Given that more than 95% of the survey respondents viewed the availability of coordinated and demographically corrected norms as moderately (18%) or very important (77%), the norms provided for the NAB represent a unique and critical feature.

Provision of Equivalent/Parallel Forms

An important aspect of neuropsychological assessment is the ability to monitor and document changes in functioning over time (Friedes, 1985; Matarazzo, Carmody, & Jacobs, 1980). In fact, the results of the survey indicated that 96% of all respondents viewed the detection of change over time as a moderately (33%) or very important (63%) characteristic of a new comprehensive neuropsychological test battery. Current neuropsychological instruments are poorly equipped to meet this goal because of a lack of equivalent, "repeatable" forms (Lezak, 1995) and a limited understanding of practice effects on neuropsychological testing (Sawrie, Chelune, Naugle, & Lueders, 1996). In the NAB, these needs were addressed in two ways. First, two parallel, equivalent forms were developed for each NAB module during the initial development phases. That is, unlike many tests with parallel forms, one original form was not created first with a secondary form developed after the fact. Rather, an initial large item pool and the ratings by Advisory Council members along with the results of pilot testing made possible the two equivalent NAB forms each with a distinct set of items created simultaneously. Second, because many repeat neuropsychological testing sessions occur 6 months or more after the initial evaluation, a test–retest reliability study of the NAB was conducted based on a 6-month retest interval.

Resulting standard errors of measurement and expected practice effects (see chapter 5) help differentiate meaningful score differences from artifactual practice effects (Ivnik et al., 1999).

Increased User-Friendliness

The NAB is more user-friendly than existing instruments with respect to modularity, portability, and face validity to the examinee. Almost 90% of the survey respondents rated modularity as either moderately (29%) or very important (60%) for a new instrument. Each of the six NAB modules is "self-contained" and may be administered independently of the other modules. In addition, 76% of survey respondents rated *portability* as either moderately or very important. NAB materials are highly portable because a minimal number of manipulatives are required, and all necessary visual stimuli are integrated into a single stimulus book for each module. Moreover, the NAB Record Forms have been created to include all necessary administration instructions on the forms themselves, thus eliminating the common difficulty of having to rely on test administration manuals for instructions. The use of such manuals often leads to the examiner's awkwardly juggling multiple books and forms while administering a test, and in the case of examiners who attempt to "memorize" administration instructions, inconsistency in administration procedures. Approximately 73% of the survey respondents rated computerized administration as only slightly important or not at all important. Although this finding is initially surprising, it is understandable because even laptop computers significantly reduce portability and raise design and psychometric problems. Thus, the NAB is administered entirely by an examiner (i.e., not by computer).

Face validity, an important (Lezak, 1995) and often overlooked aspect of neuropsychological validation (Nevo, 1985), refers to whether a test appears to measure what it purports to measure to (a) the examinees who take it, (b) the administrative personnel who decide upon its use, and (c) other technically untrained observers, such as the examinee's family (Anastasi & Urbina, 1997). Tests that lack face validity are more prone to rejection by patients with brain dysfunction who are likely to be easily frustrated and fatigued. The *face validity* of the NAB was rated by the members of the Advisory Council, and items and tasks with poor face validity ratings were eliminated or modified. Although the attractiveness of test materials is not often discussed in the literature on face validity, the NAB includes modern, inviting, and colorful stimuli, materials, and artwork, including high-quality digital photography.

Ecological Validity

Ecological validity is the functional and predictive relationship between performance on a set of neuropsychological tests during a highly structured, office-based testing session and behavior in a variety of real-world settings, such as home, work, or school (Long, 1996). Franzen and Wilhelm (1996) define the *functional relationship* as the similarity of the data-collection method to the skills required in the free and open environment. They define the *predictive relationship* as the extent to which test results can predict behavior in the open environment. More than 79% of survey respondents rated ecological validity as being either moderately or highly important attributes for a new comprehensive test battery. The development of the NAB specifically emphasized ecological validity. For example, each NAB module (with the exception of Screening) includes one Daily Living test that is designed to be highly congruent with an analogous real-world (and, by definition, multidimensional) behavior (see Table 1.5).

Table 1.5
NAB Daily Living Tests

Module	Test
Attention	Driving Scenes
Language	Bill Payment
Memory	Daily Living Memory • Medication Instructions • Name/Address/Phone Number
Spatial	Map Reading
Executive Functions	Judgment