ANCDS Bulletin Board

Mary Boyle, Ph.D.

President, ANCDS, 2005–2006

Practice Guidelines for Standardized Assessment for Persons with Traumatic Brain Injury

Lyn Turkstra, Ph.D.

University of Wisconsin, Madison

Mark Ylvisaker, Ph.D.

College of Saint Rose
Albany, New York

Carl Coelho, Ph.D.

University of Connecticut, New Haven

Mary Kennedy, Ph.D.

University of Minnesota, Minneapolis

McKay Moore Sohlberg, Ph.D.

University of Oregon, Eugene

Jack Avery, M.A.

University of Minnesota, Minneapolis

Kathryn Yorkston, Ph.D.

University of Washington, Seattle

On Behalf of the Academy of Neurologic Communication Disorders Evidence-Based Practice Project

---

This article is one of a series of publications by the Academy of Neurologic Communication Disorders and Sciences (ANCDS) working groups on evidence-based clinical practice (EBP) in neurologic communication disorders. The EPB project was initiated in 1997, when ANCDS established committees of experts to develop EBP guidelines for the following areas: dysarthria, aphasia, dementia, apraxia, and cognitive-communication disorders associated with traumatic brain injury (TBI). The scope and mission of the

Journal of Medical Speech-Language Pathology
Volume 13, Number 2, pp. ix–xxxviii

EBP project are described in detail in previous publications (Golper et al., 2001; Kennedy et al., 2002; Sohlberg et al., 2003). This article was generated by the subcommittee on cognitive-communication disorders associated with TBI, and its purpose is to examine the evidence for the use of standardized, norm-referenced tests. Evaluation and assessment using nonstandardized tests and other approaches will be addressed in a separate publication. The full Technical Report is available at www.ancds.org.

OVERVIEW OF THE GUIDELINES PROCESS

The first step in the development of EBP guidelines for the use of standardized, norm-referenced tests (referred to hereafter as *standardized tests*) was to define the target population and the terms used. Next, the committee identified key questions to be answered then established the process by which information would be obtained and the criteria for generating guidelines and standards. The process and criteria were applied, and the results were summarized in a series of guidelines for evaluation and assessment.

Target Population

The target population members were children, adolescents, and adults with cognitively based communication disorders associated with TBI. TBI was defined as an acquired injury to the brain due to applied force, and it included closed head injury (e.g., via gravitational force) and open head injury (e.g., with penetration of the skull by a missile). The committee recognized that individuals with acquired cognitive-communication disorders of nontraumatic origin (e.g., strokes in the anterior cerebral artery distribution, primary hypoxic injury due to cardiac arrest) were likely to resemble persons with TBI to the extent that the two groups had similar neuropathological findings. Thus, it was hoped that the guidelines described here would have broader application to individuals with cognitive-communication disorders due to nontraumatic, acquired lesions of structures commonly injured in TBI, such as the prefrontal cortex, corpus callosum, and mesial temporal structures.

Definition of Terms

Cognitive-communication disorder. Gillis, Pierce, and McHenry (1996) discussed the challenge of how best to label the communication disorder—or disorders—typically associated with TBI. They noted that terms used in the past have included "nonaphasic language impairment, subclinical aphasia, cognitive-linguistic impairment, cognitive-language disorders, and cognitive-communicative impairment/disorders" (p. 91). The confusion in terminology has at its roots the theoretical controversy regarding the relationship of language to other cognitive functions (Gillis et al., 1996) and may be a natural consequence of moving beyond a primary focus on linguistic functions to a broader focus on communication—including aspects of communication such as pragmatics that are heavily influenced by nonlinguistic cognitive functions. For the purposes of this review, the term "cognitive-communication disorders" was used, as defined by the American Speech-Language Hearing Association (ASHA):

> Cognitive-communication disorders encompass difficulty with any aspect of communication that is affected by disruption of cognition. Communication includes listening, speaking, gesturing, reading, and writing in all domains of language (phonologic, morphologic, syntactic, semantic, and pragmatic). Cognition includes cognitive processes and systems (e.g., attention, memory, organization, executive functions). Areas of function affected by cognitive impairments include behavioral self-regulation, social interaction, activities of daily living, learning and academic performance, and vocational performance. (ASHA, 2004, p. 2)

Evaluation and assessment. Evaluation may be defined as the act of considering or examining something in order to judge its value, quality, importance, extent or condition (Webster, 1996, p. 462). Assessment is a judgment about something based on an understanding of the situation (Webster, 1996, p. 82). That is, evaluation refers to the *process* of examination, and assessment to its *result*, although the terms are used interchangeably and are listed as synonyms for each other in some dictionaries. Standardized tests address both evaluation and assessment. That is, they provide explicit procedures for evaluating an individual's strengths and limitations and a criterion against which the clinician may make a statement of assessment, such

as "within normal limits" or "in the impaired range." Some standardized tests also permit the evaluation of contributing factors and make explicit statements regarding implications for functional communication, although in most cases neither of these is based on empirical data.

Clinical assessments can be designed to serve a variety of distinct purposes, each of which may impose unique standards and demands on the assessment process. These purposes include support of medical or neurological diagnoses, diagnostic classification by communication disorder type, acquisition of services, prognosis, generation of epidemiologic or other research data, planning intervention and supports, testing intervention hypotheses, measuring progress, and preparation for legal testimony. Thus, it is to be expected that a variety of tools are used by clinicians to achieve this variety of assessment goals.

The World Health Organization (WHO) Classification of Health Outcomes (2001) offers an alternative framework for organizing a discussion of assessment procedures. According to the most recent version of the WHO framework, the outcomes of illness and injury are considered at two main levels: *body structure and function* (impairment), and *activity/participation* (the real-world consequences of the impairment, previously referred to as disability and handicap). A third component of the model is *context*, including both environmental and attitudinal contextual factors that can serve as facilitators or barriers to effective functional activity and participation. Within each of these three components, assessment procedures can be either static or dynamic (i.e., flexible hypothesis testing, possibly with task modifications, feedback, coaching, or other forms of support). The resulting six general categories of assessment are outlined in Table 1. As most standardized assessment tools were designed for static assessment at the level of impairment and to a lesser extent at the level of activity/participation, these types of assessment are the focus of this article. The remaining types of assessment will be addressed by the Writing Committee in a subsequent article on nonstandardized assessment.

Questions To Be Answered

The committee focused on two key questions for speech-language pathologists:

1. What tests can or should the speech-language pathologist use for the evaluation and assessment of communication ability in persons with TBI?

TABLE 1. World Health Organization (WHO) Classification of Health Outcomes (Organization, 2001) and implications for evaluation and assessment.

WHO Category	Assessment Tools and Procedures
Impairment: Static	Standardized tests used to identify underlying neuropsychological and neurolinguistic strengths and weaknesses
Impairment: Dynamic	Systematic manipulation of variables to isolate the neuropsychological factors that influence success and failure on standardized tasks (i.e., "process assessment")
Activity/Participation: Static	Standardized or nonstandardized observation of the individual performing functional activities; informant data, using standardized or nonstandardized procedures (e.g., rating scales); identification of successful or unsuccessful participation in real-world domains of functioning
Activity/Participation: Dynamic	Systematic exploration of the factors that influence the individual's performance of everyday activities, including possible compensatory strategies.
Context/Environmental: Static	Systematic documentation of the cognitive and communication demands of relevant real-world contexts (i.e., "curriculum-based assessment"); evaluation of the communication and support competencies of relevant everyday people in the individual's life
Context/Environmental: Dynamic	Systematic manipulation of environmental factors, including the support behaviors of communication partners, to identify context-specific procedures for effectively supporting the individual's participation and activity

The evaluation and assessment of cognitive-communication disorders is a relatively recent addition to the scope of practice of speech-language pathology, recognized by ASHA in 1988. The advent of tests specifically designed for persons with TBI is also recent. Thus, the goal of the committee was to provide a detailed review of tests that are recommended for use with this population, including their psychometric properties and research support, to serve as a resource for practicing clinicians.

2. What is the speech-language pathologist's unique contribution to the interdisciplinary evaluation process?

Speech-language pathologists may at times question their unique contribution to the evaluation and assessment of persons with TBI. By its nature, the term cognitive-communication disorder implies that cognitive functions must be evaluated at some level. Cognitive evaluation, however, may be within the scope of practice of other team members, most notably neuropsychologists. ASHA's official position is that "speech-language pathologists play a primary role in the assessment, diagnosis, and treatment of infants, children, adolescents, and adults with cognitive-communication disorders" (ASHA, 2004, p. 3) due to their unique knowledge and skills in both cognition and communication across the lifespan. Thus, to provide a resource tailored to the needs of speech-language pathologists, based on our unique "skill set," this report focuses on cognition in the context of communication (i.e., cognitive-communication disorders).

Detailed reviews of cognitive tests are available elsewhere (e.g., Lezak, 1995; Spreen & Strauss, 1998) and cognitive functions are discussed in each of the sources of expert opinion reviewed here (see *Review of Published Expert Opinion* on p. xxv). Tests of cognitive function were included in the EBP guidelines, however, if they were recommended specifically for evaluation of communication ability by speech-language pathologists, who responded to the ANCDS survey, publishers and distributors of tests for speech-language pathologists or experts in the evaluation and assessment of communication disorders in children or adults with TBI. This report is not intended to limit the tools used by speech-language pathologists, or create practice barriers between speech-language pathologists and colleagues in other professions. Rather, the intent is to create a picture of current clinical practice and generate recommendations for the future. The latter are based on a critical evaluation of existing practices and the published literature, including the opinions of experts in the field.

Readers interested in a further discussion of issues related to scope of practice in cognitive-communication disorders are referred to two reports of the Ad Hoc Joint Committee on Interprofessional Relationships of the American Speech-Language Hearing Association and Division 40 (Clinical Neuropsychology) of the American Psychological Association. The first report (Paul-Brown & Ricker, 2003) emphasized the importance of interprofessional collaboration in testing, citing the need to avoid creating practice effects by repeated testing with the same or similar instruments and presenting conflicting messages to clients and families. The Committee advocated for greater interaction between speech-language pathologists and neuropsychologists on professional practice issues—such as the lack of appropriate standardization data for many tests currently in use—and emphasized that the two professions have complementary skill sets. The second report (Ylvisaker, Hanks, & Johnson-Greene, 2002) outlined two distinct paradigms for delivering cognitive rehabilitation services, including alternative understandings of the role of assessment in rehabilitation.

METHODS

The committee undertook the following information-gathering activities.

1. *Survey of speech-language pathologists.* A survey of practicing clinicians was posted on the National Center for Neurogenic Communication Disorders CenterNet Listserve and on the ANCDS website. The survey text is presented in Appendix A. The questions were stated broadly to elicit the maximum information from respondents. The goal of this survey was to gather information regarding current practice patterns, as a foundation for recommendations for the future.

2. *Survey of test publishers and distributors.* The committee surveyed colleagues and the exhibitors present at the 2001 annual American Speech-Language-Hearing Association convention and generated a list of test publishers and distributors who marketed standardized tests specifically for speech-language pathologists. The committee asked representatives from these companies to recommend tests for use specifically by speech-language pathologists evaluating communication ability in persons with TBI.

3. *Review of test manuals.* This review was conducted in two stages. First, all standardized tests recommended by clinicians, publishers, and distributors were reviewed to determine if TBI was included in the conceptualization, purpose, or standardization of the test. The purpose of the test and its standardization sample characteristics were summarized, and subsequent publications with supplementary data were identified. Second, for tests that were designed for persons with TBI or included individuals with TBI in their standardization sample, a detailed analysis of test reliability and validity was completed. The definitions and criteria for reliability and validity are described in Appendix B. The definitions were taken from guidelines for evaluating psychometric properties of tests, published by Anastasi and Urbina (1997). The criteria were taken from the Agency for Healthcare Research and Quality (AHRQ) Evidence-Based Practice Program (Biddle, Watson, & Hooper, 2002; www.ahrq.gov/clinic/epc/).

4. *Review of the published literature.* The committee searched for research literature related to the assessment and evaluation of communication ability after TBI, using the following databases: Medline, CINAHL, ERIC, and PsychInfo. The key word string used was "brain or head" and "injury" and "communication or language" and "assessment or test or evaluation." Articles that did not contain original data and articles in which a test was used only for the characterization of participants (i.e., the test itself was not studied) were excluded. The findings of each study were summarized.

5. *Review of published expert opinion.* For relatively new areas of professional practice such as cognitive-communication disorders, the committee determined that expert opinion was important to include as evidence. Expert opinion is typically considered to be a Class III level of evidence (Woolf, 1992).The committee members were asked to recommend published text sources of expert opinion regarding the evaluation and assessment of communication in persons with TBI. The committee reviewed each of the recommended texts and summarized the expert opinions.

RESULTS

Survey of Speech-Language Pathologists

Thirty speech-language pathologists responded to the ANCDS survey, and as a group they recommended 69 tests. Details regarding survey responses are available in the technical report. Respondents to the request for peer review of the Technical Report suggested an additional 15 tests. As the survey questions were broadly stated, respondents named tests of speech, language, pragmatic ability, and cognitive functions. For the purposes of this review, the tests of speech function were excluded from further review. The respondents were in a variety of settings, including Home Health, schools, residential care, acute care hospital units, or combined in- and out-patient rehabilitation facilities.

Respondents named both comprehensive tests and screening tests. They used the tests for a variety of purposes in addition to those specified in the test manuals, including compliance with institutional practices, screening prior to more in-depth evaluation (e.g., when only subtests of a comprehensive test were used), as a basis for determining the language level in handouts and homework assignments, and as a means to examine aspects of performance for which the test was not designed (e.g., strategy use on a list-learning or word-generation task).

Test strengths and limitations mentioned most often were related to speed of administration, scope of the test, and normative data. That is, the time required to complete test administration was reported to be an advantage for some tests and a disadvantage for others, and some tests were endorsed because of the areas they purported to test, whereas others were criticized for neglecting important aspects of communication function. The lack of relevant norms was identified as a limitation of several tests.

The survey did not ask if respondents used process-oriented or fixed battery approaches, but their responses suggested a combination of these. Several respondents used a combination of tests and subtests tailored to the needs of the patient, whereas others used protocols dictated by the institution or reimbursement sources. The survey also did not ask if respondents used static or dynamic assessment, although the responses suggested that static approaches (i.e., administration of the test as instructed and calculation of scores) were more common than dynamic approaches (i.e., flexible administration and systematic exploration of the effects of feedback, changes in task variables, and other supports).

Survey of Test Publishers and Distributors

The committee solicited test recommendations for speech-language pathologists from four test pub-

lishers (American Guidance Services, Pro-Ed, The Psychological Corporation, and Thames Valley Publishing) and four test distributors (Imaginart; Lippincott, Williams and Wilkins; Maddak, Inc.; and Super Duper Publications). The person contacted for each company was the individual recommended by the company representative (e.g., a representative at a convention or the individual responding to a telephone call to the company's order number) as the individual who would be able to inform the committee about recommendations for speech-language pathologists.

The representatives of the companies surveyed recommended 40 standardized tests. These included tests designed for TBI, as well as traditional language impairment-based tests, a test of right hemisphere function, a screening test for Alzheimer disease, tests based on a developmental language model, and neuropsychological tests of component processes such as memory, speed of processing, and attention. Although most tests were at the level of impairment, three were designed to capture communication at the level of activity or participation (WHO, 2001). The specific tests recommended are available in the technical report.

Review of Test Manuals

The recommendations of clinicians and publishers were combined, and 78 tests that were designed to capture elements of communication were selected for further review. These tests are listed in Table 2, and detailed results are available in the writing committee's technical report, available on the ANCDS website at www.ancds.duq.edu. If there were multiple versions of a test, only the most recent version was selected, in accordance with the guidelines of the American Psychological Association (2002) for best practices in assessment.

For each test, the writing committee identified the purpose of the test, including its relevance to TBI, and the characteristics of the standardization sample as reported in the test manual. The latter was obtained to provide information regarding the applicability of the standardization data to the target population of persons with TBI. In some cases, normative data were obtained from research subsequent to the publication of the test manual. The citations of those research studies are included in the technical report.

For the second stage of the test review, the committee selected only those tests that explicitly referred to TBI in the test conceptualization, purpose, or standardization sample. Thirty-one tests met these criteria, and they are described in detail in the technical report,. The reliability and validity of each test were compared against standards set by the Agency for Health Care Policy Research. These standards are listed at www.ahrq.gov/clinic/epc/ and summarized in Appendix B. In brief, the strict reliability criterion was a correlation of .90, and the relaxed criterion was a correlation of .80 between measures (e.g., Time 1 to Time 2, or Rater 1 to Rater 2). Construct reliability criteria addressed test elements such as theoretical foundation, factor structure, overlap with similar measures, and relation to outcomes beyond the test setting.

A review of the manual information revealed that the strengths of most tests were in content and face validity. All tests appeared to have been well thought out by experts, typically based on the literature relating to the construct of interest. The background and test models for each instrument were described clearly in the test manuals and were consistent with the test purpose(s).

Other elements of reliability and validity were substantially weaker across tests. The results for individual tests are summarized in Table 3, and the following is an overall summary:

- *Interrater Reliability:* 8 of 31 tests met the strict criterion for interrater reliability, 5 met either strict or relaxed criteria for most or all of the derived scores, 1 met the relaxed criterion only, 10 either did not test interrater reliability (n = 9) or did not report correlation statistics (n = 1), and 1 reported interrater reliability that failed to meet the strict or relaxed criteria.
- *Internal Consistency:* 5 of 31 tests met the strict criterion for internal consistency, 10 met either strict or relaxed criteria for most or all of the derived scores, 2 met the relaxed criterion only, and 9 either did not test internal consistency (n = 8) or did not report correlation statistics (n = 1).
- *Test-Retest Reliability:* 3 of 31 tests met the strict criterion for test-retest reliability, 10 met either strict or relaxed criteria for most or all of the derived scores, 6 met either criterion for only a few scores (e.g., the total score but not subtest scores), 2 did not meet either criterion, and 5 either did not test test-retest reliability (n = 4) or did not report correlation statistics (n = 1). It should be noted that several test authors stated that test-retest reliability was difficult to address in an evolving population (e.g., individuals in the first few weeks after TBI).

TABLE 2. List of tests recommended by clinicians, publishers, and distributors for evaluation of individuals with cognitive-communication disorders.

Alzheimer's Quick Test
Wiig, Nielsen, Minthon, &Warkentin (2002)
The Psychological Corporation

ASHA Functional Assessment of Communication Skills in Adults
Frattali, Thompson, Holland, Wohl, & Ferketic (1995)
American Speech-Language-Hearing Association

Aphasia Diagnostic Profiles
Helm-Estabrooks (1992)
Applied Symbolix

Assessment of Language Related Functional Activities
Baines, Marting, & McMartin Heeringo (1999)
The Psychological Corporation

Attention Process Training (APT) Test
Sohlberg & Mateer (2001)
Lash & Associates Publishing/Training Inc.

Bay Area Functional Performance Evaluation (BaFPE)
Bloomer & Lang (DATE)
MADDAK, Inc.

Behavior Rating Inventory of Executive Function (BRIEF)
Gioia, Isquith, Guy, Kenworthy (2000)
Psychological Assessment Resources

Behavioral Assessment of the Dysexecutive System (BADS), including the Dex questionnaire
Wilson et al. (1996)
Northern Speech Services, Inc.

Boston Diagnostic Aphasia Examination-lll (BDAE-III)
Goodglass, Barresi, & Kaplan (2000)
The Psychological Corporation

Boston Naming Test—Second Edition (BNT)
Kaplan, Goodglass, & Weintraub (2000)
Psychological Assessment Resources

Brief Test of Head Injury (BTHI)
Helm-Estabrooks and Hotz (1991)
Riverside Publishing Company

Burns Brief Inventory of Communication and Cognition
Burns (1997)
The Psychological Corporation

California Verbal Learning Test-II (CVLT-II)
Delis, Kramer, Kaplan, & Ober (2000)
The Psychological Corporation

California Verbal Learning Test for Children (CVLT-C)
Delis, Kramer, Kaplan, & Ober (1994)
The Psychological Corporation

Clinical Evaluation of Language Fundamentals, Third Edition (CELF-3)*
Semel, Wiig, & Secord (1995)
The Psychological Corporation

Cognistat
Kiernan, Mueller, & Langston (2002)
Psychological Assessment Resources, Inc.
Originally published as Neurobehavioral Cognitive Status Exam (1988). Northern California Neurobehavioral Group. Fairfax. CA: Northern California Neurobehavioral Group

Communicative Abilities of Daily Living 2 (CADL-2)
Holland, Frattali & Fromm (1999).
The Psychological Corporation

Comprehensive Assessment of Spoken Language (CASL)
Carrow-Woolfolk (1999)
American Guidance Services

Cognitive Linguistic Quick Test (CLQT)
Helm-Estabrooks (2001)
The Psychological Corporation

Children's Orientation and Amnesia Test (COAT)
Ewing-Cobbs, L., Levin, H., Fletcher, J., Miner, M., & Eisenberg, H. (1990). The Children's Orientation and Amnesia Test: Relationship to severity of acute head injury and to recovery of memory. *Neurosurgery, 27*, 683–691.

Controlled Oral Word Association (COWA) Subtest
Multilingual Aphasia Examination
Benton, Hamsher, Rey & Sivan (1994)
The Psychological Corporation

Watson-Glaser Critical Thinking Appraisal (CTA)
Watson & Glaser (1980)
The Psychological Corporation

Delis-Kaplan Executive Function System (D-KEFS)
Delis, Kaplan, & Kramer (2001)
The Psychological Corporation

Detroit Test of Learning Aptitude—Third Edition (DTLA-3)
Hammill (1991)
Pro-Ed

Discourse Comprehension Test (DCT)
Brookshire &. Nicholas (1993)
BRK Publishers

(continues)

TABLE 2. *(continued)*

Expressive One Word Picture Vocabulary Test—Third Edition (EOWPVT)
Brownell (2000)
Academic Therapy Publications

Florida Affect Battery
Unpublished.

The Fullerton Language Test for Adolescents
Thorum (1980)
Consulting Psychologists Press, Inc.

Functional Independence Measure (FIM)
Uniform Data System for Medical Rehabilitation (1996)
Based on Forer, Granger, et al. (1987). Functional Independence Measure. Buffalo, NY: The Buffalo General Hospital, State University of New York at Buffalo

Functional Linguistic Communication Inventory (FLCI)
Bayles & Tomoeda (1994)
Canyonlands Publishing, Inc.

Galveston Orientation Amnesia Test (GOAT)
Levin, H., O'Donnell, V. &, Grossman, R. (1979). The Galveston Orientation Amnesia Test: A practical scale to assess cognition after head injury. *Journal of Nervous and Mental Disease, 167*, 65–684.

Goldman-Fristoe-Woodcock Auditory Memory Battery (Recognition Memory Subtest)
Goldman, Fristoe & Woodcock (1974)
American Guidance Service

Kagan's rating scales
Unpublished.

La Trobe Communication Questionnaire
Douglas, J., O'Flaherty, C., & Snow, P. (2000). Measuring perception of communicative ability: the development and evaluation of the La Trobe communication questionnaire. *Aphasiology 14*, 251–268.

Measure of Cognitive-Linguistic Abilities (MCLA)
Ellmo, Graser, Krchnavak, Hauck, & Calabrese (1995)
Speech Bin

Mini Inventory of Right Brain Injury (MIRBI)
Pimental and Kingsbury (1989)
Pro-Ed

Mini-Mental State Examination (MMSE)
Folstein, M, Folstein, S., & McHugh (1975)
Psychological Assessment Resources

Mt Wilga Test for Higher Level Language Functioning
Christie, Clarke & Mortensen (1986).
Unpublished manuscript.
Sydney, NSW: Mt Wilga Rehabilitation Centre

Nelson-Denny Reading Test
Brown, Fishco, & Hanna (1993)
Riverside Publishing Co.

Neurobehavioral Functioning Inventory (NFI)
Kreutzer, Seel, & Marwitz (1999)
The Psychological Corporation

Paced Auditory Serial Addition Test (PASAT)
Gronwall (1977)
Available from The Psychological Corporation (1998)

Peabody Picture Vocabulary Test (PPVT-III)
Lloyd Dunn & Leota Dunn (1997)
American Guidance Services

Pediatric Test of Brain Injury
Hotz, Helm-Estabrooks & Wolf Nelson (in preparation)

Porch Index of Communicative Ability (PICA)
Porch (1971)
Consulting Psychologists Press

Preschool Language Scale-3 (PLS-3)
Zimmerman et al. (1991)
The Psychological Corporation

The Profile of Executive Control System (Pro-Ed)
Broswell et al. (2001)
Lash & Associates Publishing

Prospective Memory Screening Test
Sohlberg & Mateer (2001)
Lash & Associates Publishing

Repeatable Battery for the Assessment of Neuropsychological Status (RBANS)
Randolph (2001)
The Psychological Corporation

Reading Comprehension Battery for Aphasia-II (RCBA-II)
LaPointe & Homer (1998)
The Psychological Corporation

Rehabilitation Institute of Chicago Evaluation of Communication Problems in Right Hemisphere Dysfunction—Revised
Halper, Cherney & Burns (1996)
Aspen Publications

Receptive One Word Picture Vocabulary Test (ROWPVT)
Gardner (1990)
Academic Therapy Publications

Rey Auditory Verbal Learning Test (RAVLT)
Rey (1958), Lezak (1976, 1983)

Rivermead Behavioral Memory Test (RBMT)
Wilson, Cockburn, & Baddeley (1985)
Western Psychological Services

Rancho Los Amigos Levels of Cognitive Functioning (LOCF)
Hagan (1979)
Ellingsworth Press LLC

(continues)

Ross Information Processing Assessment—Second Edition (RIPA-2)
Ross-Swain (1996)
Pro-Ed

Ross Information Processing Assessment—Geriatric Edition (RIPA-G)
Ross-Swain & Fogle (1996)
Pro-Ed

Scales of Cognitive Ability for Traumatic Brain Injury (SCATBI)
Adamovich & Henderson (1992)
Pro-Ed

Sklar Aphasia Scale
Sklar (Publication date unknown)
Western Psychological Services

Social Skills Rating System (SSRS)
Gresham & Elliott (1990)
American Guidance Services

Speed and Capacity of Language-Processing Test (SCOLP)
Baddeley, Emslie, & Nimmo-Smith (1992)
Thames Valley Test Company

Sydney Psychosocial Reintegration
Tate. Unpublished.

Test of Auditory-Perceptual Skills—Upper Level (TAPS-UL)
Gardner (1996)
Psychological and Educational Publishers

The Awareness of Social Inference Test (TASIT)
McDonald, Flanagan, Kinch, & Rollins (2002)
Thames Valley Test Company

Test of Adolescent and Adult Language (TOAL-3)
Hammil, Brown, Larsen, & Wiederholt (1994)
Pro-Ed

Test of Adolescent and Adult Word Finding
German (1990)
DLM Teaching Resources

Test of Everyday Attention (TEA)
Robertson, Ward, Ridgeway, & Nimmo-Smith (1994)
Thames Valley Test Company

Test of Everyday Attention for Children (TEA-Ch)
Manly, Anderson, Robertson, & Nimmo-Smith (1998)
Thames Valley Test Company

Test of Language Comprehension—Expanded Edition (TLC)
Wiig & Second (1989)
The Psychological Corporation

Test of Problem Solving
Zachman et al. (1991)
Linguisystems

The Token Test- Short Form
De Renzi, E., & Faglioni, P. (1978) Normative data and screening power of a shorterd version of the TokenTest. *Brain 14,* 41–49.

Western Aphasia Battery (WAB)
Kertesz (1982)
The Psychological Corporation

Woodcock Johnson III Tests of Cognitive Abilities
Mather & Woodcock (2001)
Riverside Publishing

The Word Test—Elementary
Barrett, Huisingh, Orman, Bowers, & LoGuidice (1990)
LinguiSystems

The Word Test—Adolescent
Zachman, Huisingh, Barrett, Orman, & Blagden (1989)
LinguiSystems

Writing Process Test
Warden & Hutchinson (1992)
Applied Symbolix

*The CELF-3 was reviewed here because there were data regarding its use in TBI. The CELF-4 was not recommended and did not include TBI in the test manual and therefore was not included here.

- *Construct Validity:* 17 of 31 tests demonstrated construct validity for the constructs described in the test purpose, typically by using factor analysis; 7 demonstrated construct validity for some elements of the test purpose and not others; 1 test did not address construct validity; and 1 test failed to demonstrate construct validity. Examples of tests that demonstrated construct validity for some purposes but not others are the Clinical Evaluation of Language Fundamentals—Third Edition (CELF-3; Semel, Wiig, & Secord, 1995) and the Ross Information Processing Assessment—Second Edition (RIPA-2; Ross-Swain, 1996). The CELF-3 was valid for identifying impairments in comprehension but not expression, and was relatively weak in its ability to discriminate between children with and without language disorders. The RIPA-2

TABLE 3. Summary of tests and criteria.

Test	Interrater Reliability	Internal Consistency	Test-retest Reliability	Content Validity	Construct Validity	Criterion: Concurrent Validity	Criterion: Predictive Validity
American Speech Language Hearing Association—Functional Assessment of Communication Skills	Overall score = R Subscales: 2/5 = S 3/5 = R Qualitative score = R Qualitative Subscales: 2/4 = R	R	S	Yes	Yes	Yes	DNT
Aphasia Diagnostic Profiles	DNT	Subtests: 7/9 = S 1/9 = R	Subtests: 1/12 = S 3/12 = R	Yes	Yes	DNT	DNT
Behavior Rating Inventory of Executive Function (Parent Report Form)	Parent vs. teacher for same child = N	Indexes = S 19/32 subscales = S 13/32 subscales = R	Teacher Index = S Teacher subscales = R Parent Index = R Parent subscales: 4/8 = R	Yes	Yes	Yes	DNT
Behavioral Assessment of the Dysexecutive Syndrome	All measures = S except for Modified Six Elements maximum time =R	DNT	N	Yes	Yes for discrimination No factor analysis	Yes	DNT
Brief Test of Head Injury	DNT	Total score and Severity score = S Subtests: 5/7 = R	1st tryout sample = R Standardization sample = N	Yes	Yes	Yes	DNT
California Verbal Learning Test—Second Edition	DNT	R	R	Yes	Yes	Yes	DNT
California Verbal Learning Test for Children	DNT	R	Age 8 (1 subscore) = S Age 16 (2 subscores) = R Remainder of sample = N	Yes	Yes	Yes	DNT

(continues)

TABLE 3. *(continued)*

Test	Interrater Reliability	Internal Consistency	Test-retest Reliability	Content Validity	Construct Validity	Criterion: Concurrent Validity	Criterion: Predictive Validity
Children's Orientation and Amnesia Test	S	DNT	DNT	Yes	Yes for age effects	Yes	Yes
Cognitive Linguistic Quick Test	Clock drawing = R Generative naming = S	DNT	Subtests: 1/10 = R Language domain = R Executive function domain = S	Yes	Yes	DNT	DNT
Communication Activities of Daily Living (Second Edition)	S	S	R	Yes	Yes for discrimination	Yes	DNT
Comprehensive Assessment of Spoken Language	DNT	Core composite = S all ages Indexes: 29/32 = S 3/32 = R Subtests: 37/252 scores across all age groups = S 57/252 scores across all age groups = R	Core composite: S all ages Index scores: Age 8–0 to 10–11: 4/5 = S, 1/5 = R Age 14–0 to 16–11: 3/3 = S	Yes	Yes	Yes	DNT
Controlled Oral Word Association Subtest	S	DNT	Normal adults = R Older adults and clinical population = N	N/A	Yes	Yes	Yes
Multilingual Aphasia Examination Clinical Evaluation of Language Fundamentals (Third Edition)	Formulated Sentences = S (age 6), R (age 11), N (age 16) Word association: S (all ages)	Total language = S (all ages) Receptive language: 4/12 age groups = R 8/12 age groups = S Expressive language: 3/12 age groups = R 9/12 age groups = S	Total language = S (2/3 age groups), R (1/3 age groups) Receptive language: 2/3 age groups = R 1/3 age groups = S	Yes	Yes for comprehension No for expression Weak for discrimination	Yes	DNT

(continues)

TABLE 3. *(continued)*

Test	Interrater Reliability	Internal Consistency	Test-retest Reliability	Content Validity	Construct Validity	Criterion: Concurrent Validity	Criterion: Predictive Validity
Discourse Comprehension Test	DNT	DNT	Participants with aphasia = R Participants with right hemisphere dysfunction = S	Yes	N	Yes	DNT
Functional Independence Measure	Total = S Categories = S Items = R	S	S	Yes	Yes	Yes	Yes
Galveston Orientation and Amnesia Test	Statistics not reported	DNT	DNT	N/A	N/A	Yes	Yes
LaTrobe Communication Questionnaire	DNT	R	No	Yes	Yes	Yes	DNT
Measure of Cognitive-Linguistic Abilities	S	DNT	DNT	Yes	DNT	DNT	DNT
Mount Wilga High Level Language Test	DNT	DNT	DNT	Yes	DNT	DNT	DNT
Paced Auditory Serial Addition Test	N/A	S	S	Yes	N/A	Yes	N
Rancho Los Amigos Levels of Cognitive Functioning	N	DNT	DNT	Yes	DNT	DNT	DNT
Repeatable Battery for the Assessment of Neuropsychological Status	R	Total scale score: 6/6 age groups = S Index scores: 3/5 index scores = R For 4/6 age groups: 2/5 index scores = R	Total scale score = R Index scores = N	Yes	Yes	Yes	DNT

(continues)

TABLE 3. *(continued)*

Test	Interrater Reliability	Internal Consistency	Test-retest Reliability	Content Validity	Construct Validity	Criterion: Concurrent Validity	Criterion: Predictive Validity
Rivermead Behavioural Memory Test	S	DNT	Screening: Version A vs. B or C = R Standard profile scores: Version A vs. B, C, or D = R	Yes	Yes	DNT	DNT
Ross Information Processing Assessment (Second Edition)	S	4/10 subtests = S 5/10 subtests = R	DNT	Yes	Yes for overall score, No for analyzing task-by-task performance	Yes	DNT
Scales of Cognitive Ability for Traumatic Brain Injury (Normed Edition)	DNT	5/5 subscales = S	Subscales: 1/5 = S	Yes	Yes	Yes	DNT
The Speed and Capacity of Language Processing Test	DNT	R	Pre-standardization subjects: Speed = S, Spot = N Standardization subjects: Speed, Spot = R	Yes	Yes for Speed, limited evidence for premorbid IQ estimate	Yes	DNT
The Token Test (Shortened Form)	DNT	DNT	DNT	Yes	Yes	Yes	DNT
The Awareness of Social Inference Test	DNT	Correlation not reported	Correlation not reported	Yes	Yes for discrimination between groups	Yes	DNT
Test of Everyday Attention for Children	DNT	DNT	Subtests: 4/9 = R Reliable change index calculated for Version A vs. B	Yes	Yes	Yes for discrimination between groups	DNT

continues

TABLE 3. *(continued)*

Test	Interrater Reliability	Internal Consistency	Test-retest Reliability	Content Validity	Construct Validity	Criterion: Concurrent Validity	Criterion: Predictive Validity
Test of Language Competence—Extended	Subtests 3 and 4 = S for Level 1 and Level 2	Level 1 Composite: 6/8 ages = S 2/8 ages = R Level 2 Composite: 3/8 ages = R	Level 1 Composite = S Level 1 subtests: 5/7 = R 1/7 = S	Yes	Yes	Yes	DNT
Western Aphasia Battery	Subtests = S	S	Aphasia quotient = S Cognitive quotient = R Subtests: 7/9 = S 1/9 = R	Yes	Yes	Yes	DNT

R = Met relaxed criterion. S = Met strict criterion. N= Did not meet strict or relaxed relaxed criterion. DNT = Did not test. N/A = Information not available in the manual.

was valid for testing overall cognitive impairment but not for analyzing task-by-task performance.

- *Concurrent Validity:* 25 of 31 tests demonstrated concurrent validity, typically by obtaining significant correlations with scores from related measures, if similar tests existed, or by low correlations with unrelated measures, if the test addressed a novel construct; 4 tests did not address concurrent validity.
- *Predictive Validity:* 4 of 31 tests demonstrated predictive validity for outcomes beyond the assessment situation (e.g., health outcome at 1 year).

In summary, most of the tests reviewed had relative strengths in content and face validity and in demonstration of relationships with tests of similar or unrelated constructs. Fewer than half of the tests clearly demonstrated construct validity for all of the purposes stated in the manual. Few tests met the strict criteria for interrater reliability, internal consistency, or test-retest reliability, and many authors did not report that they had attempted to address these factors. Overall, the results suggest that many of the tests designed for use with, and commonly administered to, persons with TBI have relatively strong content validity but are weak psychometrically, and the relationships of test scores to important measures of outcome was not established empirically at the time of this review.

Review of Published Articles on Assessment

Most of the data-based articles on standardized testing were related to tests already reviewed by the committee. Citations for those articles were added to the information for that test in the Technical Report tables (www.ancds.org). There were four exceptions:

1. a practice patterns survey conducted by Ellmo, Graser, and Calabrese (1997);
2. a practice patterns survey conducted by Duff, Proctor, and Haley (2002);
3. two practice pattern surveys by Frank and colleagues (Frank & Barrineau, 1996; Frank, Williams, & Butler, 1997); and
4. a study by Turkstra (1999), who compared the performance of adolescents with and without TBI on the CELF-3 (Semel, Wiig, & Secord, 1995).

The first two are summarized next, and the CELF-3 was added to the list of tests reviewed.

Ellmo and colleagues (1997) surveyed clinicians across the United States to determine the number of hours spent in assessment, the frequency of use of standardized tests and their use relative to the use of nonstandardized measures, the perceived adequacy of standardized tests, and the areas assessed with these instruments. Responses were considered in reference to the severity of impairment of the population served. The authors mailed surveys to 307 rehabilitation programs serving adults with TBI, chosen from the National Directory of Head Injury Rehabilitation Services, and 106 surveys were returned. Two were judged to be unacceptable for analysis, so the final number analyzed was 104.

Of the survey respondents, most (85%) worked in acute hospital settings, and the average number of hours spent in assessment was 3.1 hours. The majority of respondents worked with individuals at all levels of severity and used a combination of standardized and nonstandardized tests. The proportion of assessment time dedicated to standardized versus nonstandardized tests was relatively equal, with the exception that for patients with severe impairments standardized tests comprised an average of 30% of the total battery. Respondents were highly dissatisfied with existing measures, as evidenced by the small percentage (11–24%) that rated either standardized or nonstandardized tests as adequate. The aspects of communication that were assessed included expressive and receptive language, pragmatics, reading, writing, and cognition. Nonstandardized tests were used by large numbers of respondents in each of these areas, with the fewest nonstandardized tests used for the evaluation of cognition in individuals with mild TBI (32%) and the most used for the evaluation of pragmatics (more than 80% for all but the patients with the most severe impairments).

The authors generated lists of the 10 standardized tests used for patients at each level of impairment from "minor" to "severe." The most commonly used tests across severity categories were the Boston Diagnostic Aphasia Examination (BDAE; Goodglass, Kaplan, & Barresi, 2000), the Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 2000), and the Ross Information Processing Assessment (RIPA; Ross-Swain, 1996).

According to Ellmo and colleagues (1977), the most striking finding was "the low satisfaction rate with formal, published measures and the consis-

tency of this low satisfaction rate across all levels of severity" (p. 20). They proposed that this might not be as much a reflection of the tests themselves but of the complex and varied nature of the TBI population in regard to both the diversity of possible deficits and the range of severity. They hypothesized that the lack of satisfaction might have different causes at different levels of severity—perhaps reflecting patients' inability to participate in testing at the severe end of the continuum and the lack of sensitive measures at the mild end of the continuum. They suggested that informal measures might be required to "round out the clinical picture" (p. 20).

Ellmo and colleagues (1997) discussed the role of the speech-language pathologist in reference to cognitive versus language assessment. They noted that overall, respondents gave cognitive tests more often than language tests, and emphasized the priority of measuring linguistic function. They noted the limitations of aphasia tests for individuals with TBI and suggested that experienced clinicians might be modifying these measures to suit the population. They recognized the need for further information about how these and nonstandardized tests were being used to achieve the goal of evaluating language and communication.

Duff and colleagues (2002) surveyed speech-language pathologists in North Carolina and Illinois regarding their practice patterns for individuals with mild TBI (MTBI). The survey included questions regarding the respondents' experience, caseload, and assessment. For assessment, the authors asked if clinicians were diagnosing MTBI and if so, using what instruments. Duff and colleagues then asked respondents to rank order 26 standardized tests according to frequency of use for clients with MTBI. Respondents also were asked to add tests that were not on the list and describe any nonstandardized measures used.

Duff and colleagues (2002) mailed 450 surveys to clinicians, and 203 were returned. Of those returned, 143 were completed by speech-language pathologists who had worked with individuals with MTBI in the past 3 years. As the remaining respondents did not have current experience with MTBI, those surveys contained demographic information only. The respondents generally were experienced clinicians, with 69% reporting more than 5 years' experience working with the MTBI population. Most (52%) worked in acute hospital settings, in-patient rehabilitation (38%), or out-patient rehabilitation (45%). The overlapping percentages indicated that many clinicians worked in more than one setting. The standardized tests most frequently used were the RIPA (Ross-Swain, 1996) (71%), the BDAE (Goodglass et al., 2000) (53%), and the BNT (Kaplan et al., 2000). The authors noted that instruments known to be sensitive to MTBI—such as the Paced Auditory Serial Addition Test (Gronwall & Sampson, 1974)—were used by less than 10% of the respondents. Eighteen respondents (13%) stated that they used their own nonstandardized or informal measures. Based on the results, Duff and colleagues made the following statement:

> Speech-language pathologists using assessment tools designed and standardized for disorders other than TBI should be guarded in the interpretation of results and cognizant that individuals with MTBI may present as normal according to the assessment instrument, despite apparent deficits in daily functioning . . . Aphasia batteries possess neither the specificity nor the sensitivity for the assessment of MTBI. These instruments do not assess the cognitive deficits that are the hallmark of TBI, and they are particularly insensitive to subtle deficits found in individuals with MTBI. By using aphasia instruments to assess MTBI, the clinician will not have clinically valid information on the individual and the extent of his/her deficits. Ultimately, this may prevent detection and administration of proper information and treatment referrals. (Duff et al., 2002, p. 782)

Frank and colleagues (Frank & Barrineau, 1996; Frank et al., 1997) conducted two studies to evaluate practices in assessment of adults and children with TBI. Frank and Barrineau (1996) mailed 386 surveys to speech-language pathologists in 335 hospitals across the United States. Of these, 237 were returned. The respondents represented a wide range of employment settings, including hospitals (72.%), schools (26.6%), clinics (17.3%), private practice (11.8%), residential facilities (7.1%), and skilled nursing facilities (6.3%). The authors compiled a list of the 20 most frequently used measures for cognitive-communicative assessment of adults with TBI. The assessment instruments used most frequently were the RIPA (Ross-Swain, 1996), BDAE (Goodglass et al., 2000), BNT (Kaplan et al., 2000), Scales of Cognitive Ability for Traumatic Brain Injury (SCATBI; Adamovich & Henderson, 1992), and Western Aphasia Battery (WAB; Kertesz, 1982). In addition, most survey respondents (99.6%) reported using informal testing to supplement formal test results. The authors noted that the majority of speech-language pathologists relied on traditional test batteries designed for patients with neurological disorders other than TBI (e.g., aphasia), as was found in an earlier survey by Schwartz (1989). A large number of respondents,

however, were using instruments designed specifically for TBI. For example, 31.2% of the sample reported using the SCATBI, and 21.1% reported using the Brief Test of Head Injury (BTHI; Helm-Estabrooks & Hotz, 1990). The factors primarily influencing respondents' choices of assessment tools were ease of administration, assistance in therapeutic planning, and comprehensiveness. Interestingly, these considerations far outweighed factors such as validity, reliability, and the characteristics of the standardization sample. Finally, only 39.7% of the respondents reported that their university program had prepared them adequately for evaluating and treating individuals with TBI. The authors concluded that "further opportunities for training SLPs in cognitive communicative areas related to the TBI population should be available to increase their preparation to serve these patients" (Frank & Barrineau, 1996 p. 95)

Frank and colleagues (1997) mailed 729 surveys to speech-language pathologists in a variety of settings that offered services for children with TBI (birth to 18 years of age). Of the surveys mailed, 227 were returned. Respondents represented primarily rehabilitation centers (34%), schools (27%), and hospitals (20%), with a small number of respondents from early intervention centers (8%), skilled nursing facilities (7%), and clinics or nonprofit organizations (4%). No single test was used commonly across pediatric assessment settings; in fact, the most frequently cited test, the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997), was listed by only 11% of the respondents. The authors compiled a list of the 19 tests cited most often, including the PPVT, the Clinical Evaluation of Language Fundamentals—Revised (Semel, Wiig, & Secord, 1987), the Expressive One-Word Picture Vocabulary Test (Gardner, 1990), and the Preschool Language Scale (Zimmerman, Steiner, & Pond, 1991). The remaining tests were chosen by less than 6% of respondents. Unlike the survey results from Frank and Barrineau (1996), the characteristics of the standardization samples for each test influenced respondents' choice of assessment tools, as did the tests' validity, reliability, and comprehensiveness. Finally, the majority (98%) of respondents reported using informal measures to supplement formal test results, especially for lower-functioning patients. The authors noted that speech-language pathologists in pediatric settings appeared to use general, developmentally based assessment instruments, whereas those assessing adults reported using tests designed specifically for individuals with speech, language, or cognitive impairments. Frank and colleagues suggested that the dynamic nature of cognitive-communicative deficits at the pediatric level necessitates using instruments that are designed to consider developmental variations.

Review of Published Expert Opinion

Committee members generated a list of texts commonly used by speech-language pathologists in educational and clinical settings. The texts were by Kennedy and DeRuyter (1991), Hartley (1995) Gillis and colleagues (1996), Ylvisaker and Gioia (1998), Sohlberg and Mateer (2001), and Blosser and DePompei (2003). Each source was reviewed and recommendations for standardized test use were extracted and summarized. All of the authors discussed standardized testing within the context of an evaluation process that includes history taking, observation, functional assessment, consideration of environmental factors, and the use of nonstandardized instruments. For the purposes of this report, only those recommendations related to standardized testing are described. Because of the nature of our committee and the scope of its charge, expert opinion was restricted to texts commonly used by speech-language pathologists. Later in this section we summarize relevant findings from the extensive literature in clinical neuropsychology.

Kennedy and DeRuyter (1991)

Kennedy and DeRuyter (1991) discussed four challenges in the assessment of individuals with TBI:

1. the heterogeneity of the population, both across individuals (pre- and postmorbidly) and over time postinjury;
2. limitations in existing standardized tests, including their focus on receptive and expressive language skills—which rarely are primary areas of deficit after TBI—and the lack of standardization on persons with TBI;
3. "territorial" issues among team members due to the "overlapping nature of cognition, language, and behavioral symptoms" (p. 155); and
4. issues related to the widespread use of rating scales with this population.

In regard to the last of these, Kennedy and DeRuyter recommended that clinicians examine the purpose and target population for which the instru-

ment was designed. They noted that descriptive scales are particularly useful for individuals who are unable to participate in standardized testing, but must be scrutinized in regard to interrater reliability and content validity.

Kennedy and DeRuyter (1991) discussed standardized testing within the context of an assessment approach that combines process-oriented measures (e.g., identification of task and performance variables related to errors) and functional measures (e.g., how well the individual communicates in the environment). They stated that the literature suggests standardized tests provide baseline information, help differentiate diffuse cognitive disorganization versus focal symptoms, identify underlying processes responsible for communication breakdown, and assist in determining the entry point for intervention. They recommended standardized assessment of attention, discrimination and differentiation, temporal ordering and sequencing, organizational abilities, memory processes, analysis and synthesis, higher level reasoning and abstraction, integration, and maintenance of goal-directed activity over time. These were viewed as processes underlying communication.

Kennedy and DeRuyter (1991) acknowledged that it is "accepted practice that portions from existing standardized batteries be used" and that "standardized tests designed for children, adolescents, and adults in learning disabled and aphasia populations are commonly used" (p. 161). They warned of the potential psychometric dangers of using standardized tests on a population that differs from the normative sample, raising concerns regarding content validity, reliability, and inappropriate application of norms. They stated that, "No single test battery currently exists to address the unique cognitive language problems encountered by the TBI population" (p. 161). As a result, they concluded that clinicians have "no choice but to use standard tests in nonstandard procedures" (p. 161). The authors provided a table of recommended tests, including tests for aphasia, intelligence, adolescent language, identification of learning disabilities, and component processes such as vocabulary, problem solving, and orientation. Of 23 tests and subtests listed, only one—the Galveston Orientation and Amnesia Test (Levin, O'Donnell, & Grossman, 1979)—was designed specifically for use with persons with TBI.

Hartley (1995)

Hartley (1995), in her text on functional intervention for cognitive-communication disorders after TBI, recommended standardized testing for the evaluation of "component processes" such as memory, attention, and language. She stated that, "Standardization of scores permits the detection of intraindividual variation across areas or processing strengths and weaknesses that will influence all areas of real-life functioning" (p. 80). While acknowledging the potential limitations in ecological validity associated with decontextualized assessment, Hartley noted that information derived from such tests could be used in ecologically valid contexts. For example, test data could assist in the generation of compensatory strategies to be used in functional settings.

In reference to standardized assessment of communication, Hartley (1995) focused on language and aphasia. She noted that "no single test exists that meets the needs of individuals who do not have aphasia" and that "many clinicians find that the best approach is to select aphasia battery subtests that tap the desired areas and to supplement these with other standardized tests of language functioning" (p. 81). She recommended tests such as the Minnesota Test for Differential Diagnosis of Aphasia (Schuell, 1972), the BDAE (Goodglass et al., 2000), and the WAB (Kertesz, 1982), as well as neuropsychological tests such as the Wechsler Memory Scales, Revised (Wechsler, 1987) for components such as verbal learning and immediate recall.

Hartley commented on the lack of adult standardized tests and the resulting use of child tests (e.g., the PPVT; Dunn & Dunn, 1981) or informal measures for the evaluation of adults. She also described extra-test factors commonly associated with TBI, such as sensory and motivational changes and medication effects, that may have a negative influence on the validity of test scores.

Gillis, Pierce, and McHenry (1996)

Gillis and colleagues (1996) provided a detailed description of elements of cognition (e.g., memory and attention) that relate to communication. This description included both the theoretical constructs and the measures used in research and clinical testing of each cognitive skill. In regard to language, they summarized the research on aphasia and TBI, concluding that while aphasia is not a predominant feature of TBI, many individuals with TBI—particularly in the acute stage—show signs of aphasia such as impaired repetition. In those cases, Gillis and colleagues advocated the use of behavioral descriptors rather than labels such as "nonfluent aphasia." In language evaluation, Gillis and colleagues noted the importance of recognizing

that most tests measure more than one skill, so that, for example, a reading test demands attention and memory in addition to language. Other general guidelines also were provided, including elements to include in the evaluation process and recognition of the need to consider extra-test factors such as anxiety and arousal.

In regard to other aspects of communication, based on a review of the literature Gillis and colleagues (1996) concluded that, "Structured language (aphasia) batteries, like numerous other neuropsychological batteries, have not proven to be highly sensitive to the communication impairment of traumatic brain injury" (p. 124). As a result, the authors stated, clinicians had turned to the evaluation of discourse and pragmatics, which were more likely to capture the "interface of cognitive, linguistic, and psychosocial abilities" (p. 124). Herein lies a key issue in the evaluation of cognitive-communication disorders: by their nature, these disorders are manifest at the level of activities and participation in daily living (i.e., they are the expression of cognitive impairments in everyday communication tasks), and most standardized instruments are at the level of impairment. Perhaps, as a result, most clinicians are combining extrapolations from impairment-level cognitive tests with nonstandardized evaluation in the clinic. As Gillis and colleagues noted, this raises problems that have become "the hallmarks of communication impairment following traumatic brain injury" (p. 125): the structure of the testing situation supports cognitive function and does not assess the types of communication challenges faced by persons with TBI outside of clinical settings. Also, there are many limitations of nonstandardized measures, particularly in reference to evaluation of pragmatic abilities by clinicians who are likely to differ from their clients in significant ways (see discussion of this issue in Turkstra, McDonald, & DePompei, 2001).

Gillis and colleagues (1996) summarized the results of previous studies of communication evaluation after TBI, including a survey of clinicians reported by Nelson and Schwentor (1991). Those authors found that the Reading Comprehension Battery for Aphasia (LaPointe & Homer, 1998) was used by most respondents for evaluation of reading, followed by informal testing. Informal testing was the most common method for evaluating writing, followed by the BDAE (Goodglass et al., 2000).

Ylvisaker and Gioia (1998)

Ylvisaker and Gioia (1998) focused on the evaluation of children and adolescents with TBI. Consistent with the texts just reviewed, they attributed most language and communication disorders after TBI to impairments in executive function, slow and inefficient processing of information, and problems with new explicit learning. They noted that performance on standardized language tests often is average or above average. Because of the potential insensitivity of existing standardized measures and their inherent limitations in guiding intervention planning, Ylvisaker and Gioia proposed that contextualized hypothesis testing is most critical to assessment for purposes of deriving an intervention and support plan. Hypothesis-testing assessment will be discussed in a subsequent publication on nonstandardized assessment by this committee.

Ylvisaker and Gioia (1998) identified six main types of language and communication disability after TBI, and the potentially useful evaluation procedures associated with each:

1. *Disorganized language or impaired discourse.* The authors recommended nonstandardized evaluation of discourse, using tasks such as writing a summary of a classroom lecture and manipulating variables such as length, level of abstraction, and use of compensatory strategies.
2. *Inflexibility.* This may be manifest in difficulty with ambiguous meanings and fluency in conversation and may be evaluated with tests such as the Test of Language Competence (TLC; Wiig & Secord, 1989), as well as informal measures.
3. *Concrete thinking.* Difficulty with abstraction and inference are common sequelae of injury to the prefrontal cortex and may be evaluated with metaphor or figure of speech. Measures of inference and abstraction are included in standardized tests such as The Word Test—Adolescent (Zachman, Huisingh, Barrett, Orman, & Blagden, 1989), the TLC (Wiig & Secord, 1989), and the Comprehensive Assessment of Spoken Language (Carrow-Woolfolk, 1999).
4. *Inefficient word retrieval.* To evaluate word retrieval, Ylvisaker and Gioia (1998) recommended the use of verbal fluency measures such as the SFA task (Gaddes & Crocket, 1975), the category-naming test from the CELF (Semel et al., 1987), the Test of Word Finding (German, 1989), the Test of Word Finding in Discourse (German, 1991), or the Test of Adolescent and Adult Word Finding (German, 1990). Informal evaluation with time pressure and other stressors added also was recommended.
5. *Impaired social communication.* The authors recommended using nonstandardized measures of conversational competence, such as the Pragmat-

ic Protocol (Prutting & Kirschner, 1987), with data obtained in various communication contexts relevant to the individual. They also noted that social skills inventories may be helpful.

6. *Inefficient learning of new language.* The authors recommended the use of tests of verbal memory and learning, such as the California Verbal Learning Test—Children's Version (Delis, Kramer, Kaplan, & Ober, 1994), and the Test of Memory and Learning (Reynolds & Bigler, 1993). Tests that mimic academic learning also may be used, such as the Visual-Auditory Subtest of the Woodcock-Johnson Psychoeducational Battery, Cognitive Tests (Woodcock & Johnson, 1989). The authors emphasized, however, that contextualized teaching of concepts and associated language was the most revealing evaluation procedure.

Sohlberg and Mateer (2001)

Sohlberg and Mateer (2001) discussed the evaluation of communication ability within the interdisciplinary framework of assessing cognitive function in rehabilitation. These authors presented detailed information regarding standardized testing of component processes of attention and concentration, memory and new learning, executive functions, and behavior/adjustment. They discussed general psychometric considerations in using standardized tests (e.g., reliability, validity, and internal consistency), and contrasted the fixed-battery type of neuropsychological testing approach with the process-oriented or flexible approach advocated by neuropsychologists such as Kaplan (1988) and Lezak (1995). The authors discussed functional rating scales and their use in rehabilitation, clarifying the distinction between "functional" and "nonstandardized" instruments (i.e., "functional" typically refers to the level of assessment, whereas "standardized" refers to the structure of the test, so there are standardized functional measures and nonstandardized impairment-level measures). Measures of language and communication per se were not discussed, and no specific tests were recommended. The authors concluded with guidelines for the assessment of individuals with cognitive impairments, including who should administer the tests, how to choose the most appropriate test, and elements to include in the evaluation process (e.g., both impairment- and activity/participation-level measures).

Blosser and DePompei (2003)

Blosser and DePompei (2003) discussed evaluation and assessment within the context of planning transitions for youth with TBI, including transitions from hospital to school, and from school to independent living or employment. Within this framework, the focus of evaluation is on skills the child will need in the transitional setting, and the skills to be evaluated are those that the child will need to function in that specific context. The structure of the evaluation should take into account the child's goals, interests, and strengths. The child's needs for the future should be anticipated and planned for to maximize his or her opportunity for success and accommodate growth and development. Blosser and DePompei stated that quality of life issues play a critical role in this process and that all evaluation procedures should be measured against their contribution to the child's participation in his or her community, well-being, and personal fulfillment. They noted that a one-time diagnostic evaluation may not be "the most reliable means of understanding cognitive-communication function" (p. 97), given that the child or adolescent is constantly changing, and that assessment "may be better viewed as an ongoing, interactive process" (p. 97).

Blosser and DePompei (2003) discussed standardized testing as a mechanism for understanding "all variables that contribute to the behavior that should be stimulated or altered" (p. 100). They noted that standardized testing (a) can provide information about specific aspects of cognitive function such as language comprehension, (b) may contribute to an understanding of underlying impairments in basic cognitive processes, and (c) may be necessary for the child or adolescent to qualify for services or placement. Such tests also serve as a vehicle for the collection of efficacy data, which is increasingly in demand in rehabilitation and special education. The authors listed several benefits of standardized testing, including the following:

- There is consistency in data collection across examinees.
- Most children are familiar with the tests and therefore comfortable with them.
- Tests permit the quantification of information and comparison of performance over multiple tests.
- The results are more universally accepted as indices of efficacy.
- Identification of underlying impairments may advance understanding of multiple signs and symptoms of TBI.

Blosser and DePompei (2003) described three tests that were standardized specifically for chil-

dren or adolescents with TBI: the SCATBI (Adamovich & Henderson, 1992), which may be used for individuals age 15 years or older; the Pediatric Test of Brain Injury (PTBI; Hotz, Helm-Estabrooks, & Wolf-Nelson, in preparation), which focuses on cognitive-communication functions in children; and the Clinical Performance Measure for Pediatric Brain Injury (CPM-PBI; Smith, Haley, & Coster, in press). The CPM-PBI has been retitled the Functioning After Brain Injury Scales (K. Smith, personal communication, July 2, 2004) and is expected to be available to the public in late 2005. It includes items related to the generalization of new learning, following commands, and using single words and gestures. Further information is available at www.neri.org, by choosing "Products," then "Research Instruments." The SCATBI and PTBI are included in the tests reviewed in the technical report. It should be noted that the PTBI was in development and standardization data were not available at the time of writing of this report, so the test review should be considered preliminary.

Blosser and DePompei (2003) recommended using an "individualized, selected battery of tests when determining the cognitive-communicative strengths and needs of a particular child/adolescent" (p. 102). They stated that adult tests are not valid for use with children. The authors discussed test modifications that may be revealing (e.g., allowing untimed testing), acknowledging that any change in the standardized test protocol must be documented.

The major concerns about standardized tests raised by Blosser and DePompei (2003) were summarized from their review of the literature. They included concerns about the failure of such tests to reveal impairments that are evident in less structured and more distracting environments (i.e., in the contexts in which the behaviors will be used); the failure of children to qualify for services because scores on tests of old knowledge are within normal limits; and the tests' reliance on automatic behaviors that may be preserved after injury. They also noted the potential effects of the examiner's behavior, which may support the child's performance. These concerns indicated that sources of information beyond standardized tests must be considered in the evaluation and assessment process.

Ecological Validity: A Brief Review of the Neuropsychological Literature

Concerns about the shortcomings of standardized, office-bound testing described in these communication disorders texts are consistent with recent findings in clinical neuropsychology. The ecological validity of neuropsychological assessment tools has come under increasing scrutiny. Ecological validity refers to the power of a measure to accurately predict or represent the individual's behavior in everyday life.

Historically, standardized tests have been plagued by ecological validity questions when used with individuals with frontal lobe injury and associated executive system impairment (Eslinger, Biddle, & Grattan, 1997; Wilson, 1993). Lezak (1982) characterized standardized testing of executive functions as a paradox, in that the nature of the testing context, the organized structure of the tests, and the helpful behavior of the evaluator largely eliminate the need for executive self-regulation of behavior. In everyday life, executive functions are responsible for complex behaviors that are difficult to replicate under standardized testing conditions, including self-directed goal setting; planning and organizing complex, goal-directed behavior; ongoing monitoring and evaluating of performance; flexible shifting in strategy use; and sustaining focus over time in complex, real-world contexts. Further, social disability is a major feature of frontal lobe injury, and to date, this is addressed inadequately in standardized tests (Rath, Simon, Langenbahn, Sherr, & Diller, 2000).

Because of these shortcomings of standardized assessment, Stuss and others have referred to the testing situation as the individual's "prosthetic frontal lobes" (Stuss & Benson, 1986; Stuss & Buckle, 1992). Indeed, Anderson and colleagues concluded that a disparity between test results and real-world performance is sufficiently common among children with frontal lobe lesions to be considered a diagnostic symptom of such injury: "The dissociation between severe dysfunction in daily activities and good performances on standardized cognitive tests provides both an important diagnostic indicator as well as a major challenge in the evaluation of persons with prefrontal dysfunction" (Anderson, Damasio, Tranel, & Damasio, 2000, p. 289).

As damage to the frontal lobes and associated executive function impairments are common in both children (Levin, Goldstein, Williams, & Eisenberg, 1991) and adults (Anderson, Bigler, & Blotter, 1995) with TBI, the test versus daily living context disparity is particularly critical for this population. For example, a study of adults with TBI found situational vocational assessment to be superior to a battery of neuropsychological tests in predicting vocational outcome (LeBlanc, Hayden, & Paulman,

2000). Similarly, Mangeot, Armstrong, Colvin, Yeates, and Taylor (2003) found that the Behavior Rating Inventory of Executive Functions (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000) predicted adaptive functioning and behavioral adjustment in a sample of children 5 years post TBI, whereas a battery of neuropsychological tests failed to do so. The BRIEF, which includes both teacher and parent forms, is a standardized rating scale designed to reliably obtain information from real-world informants about the child's everyday executive, self-regulatory functioning.

Concerns regarding ecological validity are by no means limited to tests of executive function. For example, quiet testing rooms can mask attention deficits, and memory tests may fail to replicate the learning contexts encountered in daily living. Several new assessment tools have been developed by clinical neuropsychologists in recent years to address concerns about the ecological validity of most standardized, office-bound assessment procedures when used with individuals with TBI (see review by Chaytor and Schmitter-Edgecombe, 2003). Few of these measures are in common clinical use, however, and studies of their validity have only begun to emerge. Furthermore, by their nature, standardized tests do not suggest the appropriate mix of impairment-oriented, activity-oriented, and context/support-oriented interventions.

Summary of Expert Opinion

The general consensus of experts in cognitive-communication disorders after TBI is that standardized tests are one component of an evaluative process that should include multiple sources of information. With notable exceptions (e.g., the ASHA-FACS, Frattali, Thompson, Holland, Wohl, & Ferketic, 1995; and the CADL-2, Holland, Frattali, & Fromm, 1999), most standardized tests are designed at the level of impairments in cognitive and linguistic functions, and as such they are useful for the identification of strengths and weaknesses that may influence performance in communication activities and participation in community and family life. Knowledge of impairment-level strengths and limitations also is an important factor in determining treatment techniques, so that, for example, the strategies to be used are within the individual's abilities. In general, however, there is a mismatch between the focus of most standardized tests, which measure outcome at the impairment level, and the signs and symptoms of cognitive-communication disorders, which are expressed at the level of communication activity and life participation.

When standardized tests are used to evaluate communication difficulties, most experts recommend a combination of cognitive tests and language or aphasia tests. Issues regarding the psychometric properties of such tests were raised by several of the experts reviewed here—including concerns regarding the validity of tests designed for other populations (e.g., the use of pediatric developmental language tests with adults or the use of aphasia tests for individuals with cognitive-communication disorders). The experts emphasized that the imposed structure of the test situation, and its lack of resemblance to the daily communication challenges faced by persons with TBI, may create a misleading picture of the person's performance outside of clinical settings.

RECOMMENDATIONS

Typically, practice recommendations are stated as standards, guidelines, or options according to quality of the supporting evidence. In the case of standardized assessment, there is a lack of empirical data to support the use of many standardized, norm-referenced tests for individuals with cognitive-communication disorders after TBI. Thus, the committee's recommendations will be summarized as *practice options* that are based on the detailed review of test manuals, relevant published studies, and the published views of experts. It must be noted, however, that as with all evaluations and assessments, clinicians must clarify the purpose of assessment before selecting tools to achieve that purpose. Even tests that meet most psychometric criteria may not serve all assessment purposes, such as designing specific intervention and support plans, and monitoring progress toward specific objectives.

Option 1

Clinicians should use caution in using most published standardized, norm-referenced tests for the evaluation and assessment of persons with TBI. The evidence reviewed here suggests there are limitations on the use of many published language and communication tests with persons with TBI. Most of the tests recommended by publishers, distributors, and clinicians were not developed for persons with TBI and did not address TBI in their development or standardization. Of the 31 tests that explicitly referred to TBI, most had clearly stated the

theoretical and clinical motivation and attempted to address reliability and validity. Most, however, did not meet established criteria for reliability and validity. Predictive and ecological validity were weaknesses. For example, only three of the tests or measures, the Galveston Orientation and Amnesia Test (Levin et al., 1979), the BRIEF (Gioia et al., 2000), and the Children's Orientation and Amnesia Test (Ewing-Cobbs, Levin, Fletcher, Miner, & Eisenberg, 1990) formally studied the relationship of scores to measures of outcome beyond test settings.

Option 2

For appropriately selected assessment purposes and age groups, the writing committee identified several standardized, norm-referenced tests that met the majority of the stated criteria for reliability, validity, and inclusion of individuals with TBI in the test design and standardization sample. The tests are the following:

- American Speech Language Hearing Association Functional Assessment of Communication Skills in Adults (ASHA-FACS; Frattali et al., 1995). The ASHA-FACS is a caregiver interview questionnaire that was developed for the purpose of measuring performance at the level of activities in adults with communication disorders. The measure was developed for use with persons with TBI, and reliability and validity were considered carefully in its design. The ASHA-FACS was normed on a North American sample and was not designed to consider individual and cultural contexts. This was illustrated in recent study by Larkin and colleagues (Larkins, Worrall, & Hickson, 2004) who found that although the ASHA-FACS items were relevant to stakeholders in New Zealand, the wording differed substantially and there were many additional items unique to the New Zealand indigenous peoples.
- Behavior Rating Inventory of Executive Function (BRIEF; Gioia et al., 2000). The BRIEF is a questionnaire designed to evaluate executive function in daily living in children. Like the ASHA-FACS, it is aimed at the level of activities (e.g., "I act wilder or sillier than my friends."). The original BRIEF was designed for school-age children. A self-report version for adolescents recently was published, and a self-report version for adults is in development.
- Communication Activities of Daily Living, Second Edition (CADL-2; Holland et al., 1999). The CADL-2 also was designed to capture performance at the activity level, in adults with communication disorders. It has a role-playing format in which the examinee is asked to generate responses related to specific scenarios from daily living.
- Functional Independence Measure (FIM; Uniform Data System for Medical Rehabilitation, 1996). The FIM was designed to meet the need for a uniform measure of disability and outcome from rehabilitation. Its relation to outcome has been studied in more than 27,000 patients across North America. Although the results of many studies have demonstrated the reliability and validity of the FIM, the items for evaluation of communication are very limited and do not capture clinically and personally meaningful improvements within its relatively insensitive 7-point rating scales.
- Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, 2001). The RBANS was designed to be used as a screening test by a variety of health professionals, including speech-language pathologists, for evaluation of cognitive impairments in adults. The test is brief and has two forms, which makes it useful for repeated evaluations of individuals with low tolerance for testing.
- Test of Language Competence—Extended (TLC-E; Wiig & Secord, 1989). The TLC-E is one of the few available standardized tests of pragmatic communication, and measures communication at the level of impairment in children and adolescents. It was found to be sensitive to communication challenges in TBI (e.g., Towne & Entwisle, 1993), although several of the stimulus items require updating to reflect current language use.
- Western Aphasia Battery (WAB; Kertesz, 1982). The WAB was included in the detailed test review because of its inclusion of cognitive measures to supplement the evaluation of language. It was designed for the evaluation of language in adults with acquired neurological disorders, and is aimed at the level of impairments in body function. As has been noted by many authors,

individuals with cognitive-communication disorders may perform poorly on language tasks such as those included in the WAB, but for reasons other than impairments in the target constructs.

Data for individuals with TBI are included for all of these tests except the RBANS and WAB. The TLC-E data are in a later publication (Towne & Entwisle, 1993). It is noteworthy that these tests represent a range of depth, from screening to comprehensive evaluation, and both impairment and activity/participation levels of the World Health Organization Classification of Health Outcomes (WHO, 2001).

The omission of many well-known cognitive tests from this list should not be interpreted as a recommendation against the use of cognitive tests by speech-language pathologists for evaluation of individuals with cognitive-communication disorders. Inspection of the technical report will reveal that many cognitive tests were reviewed and did not meet the designated criteria for use with individuals with TBI. This may have been because individuals with TBI were not included in the test's original purpose or standardization or the tests did not meet the criteria for reliability or validity. In particular, the predictive and ecological validity of many cognitive tests have not been established. The committee further recognizes that the tests discussed here were limited to those cited by survey respondents, which was a relatively small sample, and test publishers and distributors. Although the respondents recommended many tests, the list likely would be expanded with a larger survey sample.

The committee recognizes that evaluation of cognitive functions underlying communication ability is within the scope of practice of speech-language pathologists. There are several comprehensive reviews of neuropsychological tests that speech-language pathologists may consult for further information (e.g., Lezak, 1995; Spreen & Strauss, 1998) when information about a particular construct is relevant to the goals of intervention. In this context, it is critical that speech-language pathologists obtain appropriate training in the administration of cognitive tests, as with any assessment tool. It also is important to recognize that many cognitive tests are not available to speech-language pathologists because they do not meet the qualification requirements established by the test publisher (e.g., The Psychological Corporation).

Option 3

The writing committee recommends, in accordance with the consensus of expert opinion, that standardized testing must be considered within a broader framework that considers evaluation of the person's pre-injury characteristics, stage of development and recovery, communication-related demands of personally meaningful everyday activities and life and communication contexts. It is recognized that standardized, norm-referenced tests may have limited ecological and predictive validity beyond test settings, particularly tests that measure performance at the level of body structure and function (i.e., impairment), although they may play an important role in delineating strengths and weaknesses relevant to treatment planning. Standardized tests may assist clinicians in identifying the underlying cognitive impairments associated with more realistic and relevant everyday activities. Decisions about intervention and support, however, require supplementary, dynamic, nonstandardized evaluation procedures and evaluation of the communication context. These dimensions of comprehensive assessment will be addressed in a subsequent publication by the writing committee.

Option 4

In accordance with the report of the Ad Hoc Joint Committee on Interprofessional Relationships of the American Speech-Language Hearing Association and Division 40 (Clinical Neuropsychology) of the American Psychological Association (Paul-Brown & Ricker, 2003; Ylvisaker et al., 2002), the writing committee recommends that speech language pathologists integrate their cognitive-communication assessments with those of other professionals whose scope of practice includes cognitive assessment, most notably neuropsychology. Many practitioners appear to be using parts of neuropsychological tests in their evaluations, particularly in time-constrained settings, and may not recognize the significant limitations in interpreting these tests. From a test development perspective, practitioners in speech-language pathology benefit from greater collaboration and integration with experts in related fields such as neuropsychology and educational psychology, which tend to generate instruments that are aimed at similar constructs but have been more extensively evaluated in terms of reliability and validity.

Recommendations for Future Research

The need for research in three main areas emerged from the review of evidence for evaluation and assessment in TBI.

1. There is a clear need for norm-referenced measures that meet clinicians' needs for the evaluation of individuals with cognitive-communication disorders associated with frontal lobe injury. It is widely recognized that developmental language tests and aphasia tests have limited application in this population and that highly structured impairment-based instruments are unlikely to reveal the most relevant challenges faced by persons with frontal system dysfunction. It is likewise recognized that standards of comparison are necessary in many contexts, particularly for functions such as pragmatic communication ability, which might be particularly sensitive to examiner bias in interpretation. Thus, while nonstandardized assessment plays a significant role in the evaluation process, there also is a need for standardized, norm-referenced instruments designed for the evaluation of individuals with cognitive-communication disorders and with proven reliability and validity.

2. There is a need for studies relating standardized test data to outcomes beyond clinical settings. Although tests at the level of activity/participation, such as the ASHA-FACS and CADL-2, are an important step in this direction, there remains a need to relate scores on these tests to important indices of communication quality of life, such as the Quality of Communication Life Scale (Paul et al., 2005). Furthermore, in a study by Larkins (2003), clinicians judged that individualizing the items in standardized, functional cognitive-communication scales increased their validity and clinical utility in a sample of adults with TBI.

3. There is a need for additional normative data in several populations. None of the tests reviewed identified individuals from minority populations as a distinct group in normative samples, although these individuals comprise a substantial portion of the TBI population in the United States. Older individuals currently have the highest incidence of TBI, yet there are very limited normative data for this age-group. Data for children likewise are lacking.

Introduction to the Nonstandardized Assessment Project

In a follow-up technical report and article, the TBI writing committee will expand on the theme of functional assessment, focusing on (1) the distinction between static and dynamic assessment, (2) the WHO categories of activity/participation and context, and (3) assessment for purposes of developing intervention and support plans and monitoring progress toward specific objectives. We will report on the nonstandardized instruments that survey respondents described, including screening measures of communication and cognition. We will highlight the assessment of discourse and social-communication abilities and the process of collaborative and context-sensitive hypothesis testing as a means of designing effective intervention plans.

Acknowledgments This work was supported by funding from the American Speech-Language-Hearing Association (ASHA), ASHA Division 2: Neurophysiology and Neurogenic Speech and Language Disorders, and the Department of Veterans Affairs. Support also was provided by the University of Minnesota and Case Western Reserve University. The authors wish to thank the many students and clinicians who contributed to the project, including Barbara Ambuske, Kristen Baker, Jamie Mayer, Nichole Orsini, and Kate Ruth.

Address correspondence to Lyn Turkstra, Ph.D., Department of Communicative Disorders, University of Wisconsin–Madison, 1975 Willow Drive, Madison, WI 53706 USA.
e-mail: lsturkstra@wisc.edu

BIBLIOGRAPHY

Adamovich, B., & Henderson, J. (1992). *Scales of Cognitive Ability for Traumatic Brain Injury* (normed ed.). Chicago, IL: The Riverside Publishing Company.

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57*(12), 1060–1073.

American Speech-Language-Hearing Association. (2004). Roles of speech-language pathologists in the identification, diagnosis, and treatment of individuals with cognitive-communication disorders. Practice guidelines and policies. *ASHA, Supplement 24.*

Anastasi, A., & Urbina, S. (1997). *Psychological testing.* Upper Saddle River, NJ: Prentice-Hall.

Anderson, C. V., Bigler, E. D., & Blotter, D. D. (1995). Frontal lobe lesions, diffuse damage, and neuropsychological functioning in traumatic brain-injured patients. *Journal of Clinical and Experimental Neuropsychology, 17,* 900–908.

Anderson, S. W., Damasio, H., Tranel, D., & Damasio, A. R. (2000). Long-term sequelae of prefrontal cortex damage acquired in early childhood. *Developmental Neuropsychology, 18*(3), 281–296.

Biddle, A. K., Watson, L. R., & Hooper, C. R. (2002). Using Evidence-based Research in Speech-Language Pathology. Presentation at the Academy of Neurologic Communication Disorders and Sciences Annual Meeting, Atlanta, GA.

Blosser, J. L., & DePompei, R. (2003). *Pediatric traumatic brain injury: Proactive intervention.* Clifton Park, NY: Thompson Learning.

Carrow-Woolfolk, E. (1999). *Comprehensive Assessment of Spoken Language.* Circle Pines, MN: American Guidance Service.

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review, 13*(4), 181–197.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1994). *California Verbal Learning Test–Children's Version* (CVLT-C). San Antonio, TX: The Psychological Corporation.

Duff, M. C., Proctor, A., & Haley, K. (2002). Mild traumatic brain injury: Assessment and treatment procedures used by speech-language pathologists. *Brain Injury, 16*(9), 773–787.

Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test-III* (3rd ed.). Circle Pines, MN: American Guidance Services, Inc.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test* (Revised ed.). Circle Pines, MN: American Guidance Service.

Ellmo, W. J., Graser, J. M., & Calabrese, D. B. (1997). Methods of assessment utilized by speech-language pathologists with traumatically brain injured adults: A national survey. *Journal of the New Jersey Speech and Hearing Association, 6*, 17–23.

Eslinger, P. J., Biddle, K. R., & Grattan, L. M. (1997). Cognitive and social development in children with prefrontal cortex lesions. In N. A. Krasnegor, G. R. Lyon, & P. S. Goldman-Rakic (Eds.), *Development of the prefrontal cortex: Evolution, neurobiology, and behavior* (pp. 295–235). Baltimore, MD: Paul Brookes Publishing.

Ewing-Cobbs, L., Levin, H., Fletcher, J., Miner, M., & Eisenberg, H. (1990). The Children's Orientation and Amnesia Test: Relationship to severity of acute head injury and to recovery of memory. *Neurosurgery, 27,* 683–691.

Frank, E. M., & Barrineau, S. (1996). Current speech-language assessment protocols for adults with traumatic brain injury. *Journal of Medical Speech-Language Pathology, 4*(2), 81–101.

Frank, E. M., Williams, A. R., & Butler, J. G. (1997). Current socio-cognitive communication assessment protocols for children with traumatic brain injury. *Journal of Medical Speech-Language Pathology, 5*(2), 97–112.

Frattali, C., Thompson, C., Holland, A., Wohl, C., & Ferketic, M. (1995). *American Speech Language Hearing Association Functional Assessment of Communication Skills for Adults.* Rockville, MD: American Speech Language Hearing Association.

Gaddes, W., & Crocket, D. (1975). The Spreen-Benton Aphasia Test: Normative data as a measure of normal language development. *Brain and Language, 2,* 257.

Gardner, M. (1990). *Expressive One Word Picture Vocabulary Test:* Academic Therapy Publications.

German, D. (1989). *Test of Word Finding.* Allen, TX: DLM Teaching Resources.

German, D. (1990). *Test of Adolescent/Adult Word Finding.* Allen, TX: DLM Teaching Resources.

German, D. (1991). *Test of Word Finding in Discourse.* Allen, TX: DLM Teaching Resources.

Gillis, R. J., Pierce, J. N., & McHenry, M. (1996). *Traumatic brain injury rehabilitation for speech-language pathologists.* Boston: Butterworth-Heinemann.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Function.* Odessa, FL: Psychological Assessment Resources, Inc.

Golper, L. C., Wertz, R. T., Frattali, C. M., Yorkston, K. M., Myers, P., Katz, R., et al. (2001). *Evidence-based practice guidelines for the management of communication disorders in neurologically-impaired individuals: Project introduction.* http://www.ancds.duq.edu/guidelines.html (Accessed June 1, 2004)

Goodglass, H., Kaplan, E., & Barresi, B. (2000). *Boston Diagnostic Aphasia Examination* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Gronwall, D. M. A., & Sampson, H. (1974). *The psychological effects of concussion.* New Zealand: Auckland University Press.

Hartley, L. L. (1995). *Cognitive-communicative abilities following brain injury: A functional approach.* San Diego, CA: Singular Publishing Group.

Helm-Estabrooks, N., & Hotz, G. (1990). *Brief Test of Head Injury* (Standardized Edition). Chicago, IL: The Riverside Publishing Company.

Holland, A., Frattali, C., & Fromm, D. (1999). *Communication Activities of Daily Living* (2nd ed.). Austin, TX: Pro-Ed.

Hotz, G., Helm-Estabrooks, N., & Wolf-Nelson, N. (in preparation). *Pediatric Test of Brain Injury.*

Kaplan, E. (1988). A process approach to neuropsychological assessment. In T. Boll & B. K. Bryant (Eds.), *Clinical neuropsychology and brain function: Research, measurement, and practice* (pp. 125–167). Washington, DC: American Psychological Association.

Kaplan, E., Goodglass, H., & Weintraub, S. (2000). *Boston Naming Test* (2nd ed.). San Antonio, TX: The Psychological Corporation.

Kennedy, M., Avery, J., Coelho, C., Sohlberg, M., Turkstra, L., Ylvisaker, M., et al. (2002). Evidence-based practice guidelines for cognitive-communication disorders after traumatic brain injury: Initial report of the Academy of Neurologic Communication Disorders and Sciences Writing Committee for Practice Guidelines in Cognitive-Communication Disorders after Traumatic Brain Injury. *Journal of Medical Speech-Language Pathology, 10*(2), ix–xiii.

Kennedy, M., & DeRuyter, F. (1991). Language and cognitive basis for communication disorders following traumatic brain injury. In D. Beukelman & K. Yorkston (Eds.), *Communication disorders following traumatic brain injury: Management of cognitive, language, and motor impairments* (pp. 123–190). Austin, TX: Pro-Ed.

Kertesz, A. (1982). *Western Aphasia Battery* (1st ed.). San Antonio, TX: The Psychological Corporation.

LaPointe, R., & Homer, J. (1998). *Reading Comprehension Battery for Aphasia* (2nd ed.). San Antonio, TX: The Psychological Corporation.

Larkins, B. M. (2003). *Functional communication assessment of individuals with traumatic brain injury living in New Zealand.* Unpublished dissertation, University of Queensland, Brisbane.

Larkins, B. M., Worrall, L. E., & Hickson, L. M. (2004). Stakeholder opinion of functional communication activities following traumatic brain injury. *Brain Injury, 18*(7), 691–706.

LeBlanc, J. E., Hayden, M. E., & Paulman, R. G. (2000). A comparison of neuropsychological and situational assessment for predicting employability after closed head injury. *Journal of Head Trauma Rehabilitation, 15*(4), 1022–1040.

Levin, H., O'Donnell, V., & Grossman, R. (1979). The Galveston Orientation Amnesia Test: A practical scale to assess cognition after head injury. *Journal of Nervous and Mental Disease, 167,* 675–684.

Levin, H. S., Goldstein, F. C., Williams, D. H., & Eisenberg, H. M. (1991). The contribution of frontal lobe lesions to the neurobehavioral outcome of closed head injury. In H. S. Levin, H. M. Eisenberg, & A. I. Benton (Eds.), *Frontal lobe function and dysfunction* (pp. 318–338). New York: Oxford University Press.

Lezak, M. (1995). *Neuropsychological assessment* (2nd ed.). New York: Oxford University Press.

Lezak, M. D. (1982). The problem of assessing executive functions. *International Journal of Psychology, 17,* 281–297.

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

Mangeot, S., Armstrong, K., Colvin, A. N., & Yeates, K. O., & Taylor, H. G. (2002). Long-term executive function deficits in children with traumatic brain injuries: Assessment using the Behavior Rating Inventory of Executive Functions (BRIEF). *Child Neuropsychology, 8*(4), 271–284.

Nelson, N. W., & Schwenter, D. (1991). Reading and writing disorders. In D. R. Beukelman & K. M. Yorkston (Eds.), *Communication disorders following traumatic brain injury: Management of cognitive, language, and motor impairments* (pp. 191–249). Austin, TX: Pro-Ed.

Paul, D., Frattali, C. M., Holland, A. L., Thompson, C. K., Caperton, C. J., & Slater, S. (2005). *Quality of Communication Life Scale.* Rockville, MD: American Speech-Language-Hearing Association.

Paul-Brown, D., & Ricker, J. H. (2003). Evaluating and treating communication and cognitive disorders: Approaches to referral and collaboration for speech-language pathology and clinical neuropsychology. Technical Report. *ASHA Supplement, 23,* 47–58.

Prutting, C., & Kirschner, D. (1987). A clinical appraisal of the pragmatic aspects of language. *Journal of Speech and Hearing Disorders, 52,* 105–119.

Randolph, C. (2001). *Repeatable Battery for the Assessment of Neuropsychological Status* (1st ed.). San Antonio, TX: Psychological Corporation.

Rath, J. F., Simon, D., Langenbahn, D. M., Sherr, R. L., & Diller, L. (2000). Measurement of problem-solving deficits in adults with acquired brain damage. *Journal of Head Trauma Rehabilitation, 15*(1), 724–733.

Reynolds, C., & Bigler, E. (1993). *Test of Memory and Learning.* Austin, TX: Pro-Ed.

Ross-Swain, D. (1996). *Ross Information Processing Assessment* (2nd ed.). Austin, TX: Pro-Ed.

Schuell, H. (1972). *Minnesota Test for Differential Diagnosis of Aphasia* (Revised ed.). Minneapolis: University of Minnesota Press.

Schwartz, R. (1989). Early rehabilitation in trauma centers: Have speech-language pathology services progressed? *American Speech-Language-Hearing Association, 31,* 91–94.

Semel, E., Wiig, E., & Secord, W. (1987). *Clinical Evaluation of Language Fundamentals* (Revised ed.). San Antonio, TX: The Psychological Corporation.

Semel, W., Wiig, E. H., & Secord, W. A. (1995). *Clinical Evaluation of Language Fundamentals* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Smith, K., Haley, S., & Coster, W. (in press). *Clinical Performance Measure for Pediatric Brain Injury* (CPM-PBI).

Sohlberg, M. M., Avery, J., Kennedy, M., Ylvisaker, M., Coelho, C., Turkstra, L., et al. (2003). Practice guidelines for attention training. *Journal of Medical Speech-Language Pathology, 11*(3), xix–xxxix.

Sohlberg, M. M., & Mateer, C. A. (2001). *Cognitive rehabilitation: An integrative neuropsychological approach.* New York: The Guilford Press.

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests* (2nd ed.). New York: Oxford University Press.

Stuss, D. T., & Benson, D. F. (1986). *The frontal lobes.* New York: Raven Press.

Stuss, D. T., & Buckle, L. (1992). Traumatic brain injury: Neuropsychological deficits and evaluation at different stages of recovery and in different pathologic subtypes. *Journal of Head Trauma Rehabilitation, 7,* 40–49.

Towne, R. L., & Entwisle, L. M. (1993). Metaphoric comprehension in adolescents with traumatic brain injury and in adolescents with language learning disability. *Language, Speech and Hearing Services in Schools, 24,* 100–107.

Turkstra, L. (1999). Language testing in adolescents with brain injury: A consideration of the CELF-3. *Language, Speech, and Hearing Services in Schools, 30,* 132–140.

Turkstra, L., McDonald, S., & DePompei, R. (2001). Social information processing in adolescents: Data from normally-developing adolescents and preliminary data from their peers with traumatic brain injury. *Journal of Head Trauma Rehabilitation, 16*(5), 469–483.

Uniform Data System for Medical Rehabilitation. (1996). *Functional Independence Measure.* Buffalo, NY: University at Buffalo.

Webster, J. S. (Ed.). (1996). *Webster's College Dictionary*. New York: Random House.

Wechsler, D. (1987). *Wechsler Memory Scale* (revised ed.). New York: The Psychological Corporation.

Wiig, E., & Secord, W. (1989). *Test of Language Competence-Expanded Edition.* San Antonio, TX: Psychological Corporation.

Wilson, B. A. (1993). Ecological validity of neuropsychological assessment: Do neuropsychological indices predict performance in everyday activities. *Applied and Preventive Psychology, 2,* 209–215.

Woodcock, R., & Johnson, M. (1989). *Woodcock-Johnson Psycho-Educational Test Battery* (revised ed.). Allen Park, TX: Teaching Resources.

Woolf, S. H. (1992). Practice guidelines, a new reality in medicine: II. Methods of developing guidelines. *Archives of Internal Medicine, 152,* 956–962.

World Health Organization. (2001). *International classification of functioning, disability and health* (report). Geneva: Author.

Ylvisaker, M., & Gioia, G. (1998). Cognitive assessment. In M. Ylvisaker (Ed.), *Traumatic brain injury rehabilitation* (2nd ed., pp. 159–179). Boston: Butterworth-Heinemann.

Ylvisaker, M., Hanks, R., & Johnson-Greene, D. (2002). Perspectives on rehabilitation of individuals with cognitive impairment after brain injury: Rationale for reconsideration of theoretical paradigms. *Journal of Head Trauma Rehabilitation, 17*(3), 191–209.

Zachman, L., Huisingh, R., Barrett, M., Orman, J., & Blagden, C. (1989). The Word Test—Adolescent. Illinois: LinguiSystems.

Zimmerman, I., Steiner, V., & Pond, R. (1991). *Preschool Language Scale* (3rd ed.). San Antonio, TX: The Psychological Corporation.

APPENDIX A

Survey for Practicing Speech-Language Pathologists

The Academy of Neurologic Communication Disorders and Sciences (ANCDS) has formed a committee that is responsible for developing practice guidelines for treating individuals with Traumatic Brain Injury (TBI). We are specifically interested in the assessment of individuals with TBI. Our goal is to be as thorough as possible in our review of available assessment tools that SLP's use with individuals with TBI. It is our hope that you will be able to give a few minutes of your time to inform us about the assessment tools you use and why you use them.

Thank you. Your time is greatly appreciated.

Questions:

1. What is your workplace?
   (a) Setting—e.g., Hospital, Community private practice, Nursing home
   (b) Service—e.g., Inpatient—subacute, acute, long-term; Outpatient

3. Who is on your caseload?
   (a) Typical diagnoses:
   (b) Age range:

4. What tools do you use to assess communication ability in persons with TBI?
   (a) list each tool
   (b) give your main rationale for using each tool
   (c) list strengths and limitations of each tool
   A chart is attached to help you complete this section.

5. Does your place of work utilize a site specific assessment tool? If yes, please describe the tool and the rationale for its use.

6. Would you be willing to share this site-specific assessment tool with us? Yes No
   If yes, please send to: (Data collection site address)

Thank you for taking the time to respond and assist the ANCDS committee.

APPENDIX B

Definitions and Criteria for Assessment of Reliability and Validity

Reliability

Interrater Reliability. Interrater reliability, also referred to as "inter-scorer" reliability, is the degree to which two independent raters agree on the score for a given task. The strict criteria for standardized tests are a simple correlation (r) between two ratings of ≥ .90, or a Kappa ≥ .80. The relaxed criteria are a simple correlation of ≥ .80, or a Kappa ≥ .70.

Internal Consistency. Internal consistency, also measured as "split-half reliability" is the degree to which a test measures a single construct. The strict criterion for internal consistency is a Cronbach Alpha ≥ .90, and the relaxed criterion is a Cronbach Alpha ≥ .80.

Test-Retest Reliability. Test-retest reliability is the degree to an individual is expected to achieve the same score when tested on two occasions. Differences from Time 1 to Time 2 may reflect recovery or degeneration, in an individual with an evolving neurological disorder, or extraneous variance in an individual with a stable neurological deficit or a neurologically normal individual. The strict criterion for test-retest reliability is a simple correlation of ≥ .90, and the relaxed criterion is a test-retest correlation of ≥ .80.

Validity

Content/Face Validity. Content validity reflects the degree to which a test is model-based, its items have graded difficulty, any source of systemic bias (e.g., cultural, racial) has been evaluated, and experts are involved in test design. Face validity refers to the test's superficial resemblance to the skills it purports to measure—e.g., if photographs on a test for older adults are chosen to represent their cohort knowledge, or if a test of long-term memory requires recalling a story after a period of time. Therefore, a test was considered to have met this criterion if it was based on an explicitly stated model, involved experts and a review of the literature in its development, addressed bias, and contained items with a superficial resemblance to the constructs of interest.

Construct Validity. Construct validity is the extent to which test measures a theoretical construct. It is measured by analyzing developmental changes, correlating performance with that on other tests, performing a factor analysis of responses, and evaluating internal consistency, convergent validity, and discriminant validity. Therefore, a test was considered to have met this criterion if scores showed developmental changes or changes over recovery (if appropriate), had predictable relationships with other measures of similar constructs, discriminated among clinical and nonclinical samples (if this was a test purpose), and had a factor structure that supported the test purpose.

Criterion-Related Validity: Concurrent. Concurrent criterion-related validity is the extent to which a test correlates with scores/other measures of behavior. For example, scores on a cognitive test may correlate with Glasgow Coma Scale scores, or scores on an aphasia test may be higher in individuals with smaller strokes. Therefore, a test was considered to have met this criterion if test scores were significantly related to other indices of behavior.

Criterion-Related Validity: Predictive. Predictive criterion-related validity is the extent to which a test predicts any criterion over time. For example, a test administered in a clinical setting may predict work success, or a test administered in the acute stage may predict performance on similar measures at one year postinjury. Therefore, a test was considered to have met this criterion if it predicted performance on other measures beyond the construct of interest. Note that predictive validity does not imply ecological validity. That is, the test may predict performance on other measures, but those measures may not reflect performance in real-life contexts.