8. Supplemental Aphasia Tests: Frequency of Use and Psychometric Properties

Susan T. Jackson and Connie A. Tompkins

To gain a comprehensive understanding of an aphasic person's linguistic capabilities, a battery of general language and supplemental measures is necessary (Wertz, 1985). Although a 1980 survey of 76 training programs accredited by the American Speech and Hearing Association (ASHA) yielded information about the most frequently used general language assessment tools for adults (Muma, Pierce, & Muma, 1983), there are no data regarding the use of supplemental aphasia tests.

Numerous factors go into a decision about whether or not to use a particular supplemental test with a particular aphasic person. One of these factors is the amount of useful information that is provided: useful in terms of more detailed information about a specific linguistic process than that which is gleaned from the general language measure, useful in terms of information about where in a particular linguistic process the aphasic person begins to break down, useful in terms of information about where to focus therapy. Another factor is the test's psychometric properties. A test's usefulness and its psychometric properties are related in the sense that psychometric soundness is necessary but not sufficient for a test to be useful. The psychometric properties of nine frequently used tests of general language ability and functional communication in aphasia have been reported (Skenes & McCauly, 1985). Although there are some psychometric data reported for some supplemental aphasia tests in the *Mental Measurement Yearbooks* (Conoley & Kramer, 1989), there has been no systematic, in-depth evaluation of supplemental aphasia tests' psychometric soundness.

The purpose of the present investigation was threefold: (a) to examine, by way of a survey analysis, the use of supplemental aphasia tests by clinical aphasiologists; (b) to evaluate the psychometric properties, based on information provided in the test manuals, of those supplemental

aphasia tests most frequently in use; and (c) to identify studies that reported psychometric data *after* the most frequently used tests were published.

METHOD

A list of 18 supplemental aphasia tests recommended by Brookshire (1986) and Wertz (1985) was sent to all of the 1988 and 1989 Clinical Aphasiology Conference participants. We sought responses from this particular sample of speech-language pathologists because the majority are experienced aphasia clinicians. Seventy of the 112 surveys were returned. Of these 70 respondents, 6 disqualified themselves: 2 respondents were not speech-language pathologists, 1 no longer worked with aphasic adults, 2 worked as research speech-language pathologists, and 1 was a master's degree student who had never had clinical contact with aphasic patients. Incomplete data from another 6 surveys were not tabulated. Thus, responses from 58 clinicians were tallied.

Survey respondents were asked to indicate whether they used each of the tests on the list often, occasionally, or seldom. They were not informed that a psychometric evaluation of the tests would ensue. Respondents were also encouraged to add any test that they felt was missing, and to make any other pertinent comments.

RESULTS

Respondent Characteristics

Of the 58 respondents, 20 had a master's degree, 35 had a doctoral degree, 1 had a bachelor's degree, and 2 had "other" degrees. This sample of aphasiologists is atypical of speech-language pathologists in general because it consists of so many doctoral-level clinicians. On the whole, the respondents were highly experienced clinicians (M = 12.6 years, range = 1–30 years). Just over half of the respondents spent 50% or more of their time in clinical contact with aphasic persons.

Supplemental Aphasia Tests: Frequency of Use

Table 8.1 shows the nine most frequently used supplemental aphasia tests in decreasing frequency of use. Only 12 respondents indicated that they used a supplemental aphasia test that was not listed on the questionnaire.

TABLE 8.1. THE NINE MOST FREQUENTLY USED SUPPLEMENTAL APHASIA TESTS

Tests Used Occasionally or Often	*% Respondents*
Boston Naming Test (*BNT*) (Goodglass & Kaplan, 1983)	98
Reading Comprehension Battery for Aphasia (*RCBA*) (LaPointe & Horner, 1979)	88
Word Fluency Test (Borkowski, Benton, & Spreen, 1967)	74
Coloured Progressive Matrices (*CPM*) (Raven, 1965)	71
Revised Token Test (*RTT*) (McNeil & Prescott, 1978)	62
Peabody Picture Vocabulary Test–Revised (*PPVT-R*) (Dunn & Dunn, 1981)	50
Auditory Comprehension Test for Sentences (*ACTS*) (Shewan, 1979)	40
Nelson Reading Skills Test (*NRST*) (Hanna, Schell, & Schreiner, 1977)	35
Reporter's Test	26

Psychometric Evaluation

Seven psychometric criteria proposed by the American Psychological Association (1974) were used to evaluate the nine most frequently used supplemental aphasia tests. These criteria were test development, tester qualifications, test interpretation, directions, validity (construct only), score stability, and norms. The definitions of each of these criteria are listed in Appendix 8.A. The decision rules used to assess whether each test met a particular criterion are provided in Appendix 8.B. These criteria and decision rules were the same as those used by Skenes and McCauly (1985). However, unlike Skenes and McCauly, half credit was given if a set of decision rules was only partially met.

Psychometric evaluation was based solely on the information presented in the test manual. The first author evaluated each test on each of the psychometric criteria, and four of the tests—the *Boston Naming Test* (*BNT*), the *Revised Token Test* (*RTT*), the *Nelson Reading Skills Test* (*NRST*), and the *Reading Comprehension Battery for Aphasia* (*RCBA*)—were rated by the second author for purposes of reliability. Inter-rater agreement was 93%; we disagreed on one criterion on one test.

Table 8.2 shows the results of the psychometric evaluation of the nine most frequently used supplemental aphasia tests. The data indicate that the test used most often (*BNT*) met the fewest psychometric criteria. It is also evident that the tester qualification and validity criteria were met least often.

TABLE 8.2. RESULTS OF PSYCHOMETRIC EVALUATION OF NINE SUPPLEMENTAL APHASIA TESTS

	Psychometric Criteria								
Test	TEST DEVELOPMENT	TESTER QUALIFICATIONS	INTERPRETATION	DIRECTIONS	VALIDITY	STABILITY	NORMS	*# Met in Part*	*# Met Fully*
BNT[a]				.5			.5	2	
RCBA[b]	.5		1	.5		.5		3	1
Word Fluency Test	1			.5			.5	2	1
CPM[c]	1	1	1	.5	*	*	.5	3	3
RTT[d]	.5		1	.5	1	.5	.5	4	2
PPVT-R[e]	1	1		1	1	1	1		6
ACTS[f]	1	.5	1	1	1	.5	.5	3	4
NRST[g]	1	1	1	1		1	1		6
Reporter's Test			.5	.5	1		.5	3	1

[a]*Boston Naming Test.*
[b]*Reading Comprehension Battery for Aphasia.*
[c]*Coloured Progressive Matrices.*
[d]*Revised Token Test.*
[e]*Peabody Picture Vocabulary Test–Revised.*
[f]*Auditory Comprehension Test for Sentences.*
[g]*Nelson Reading Skills Test.*
* = incomplete data.

Five tests met the *test development* criterion fully. Although the *RTT* provided good information about intra- and intersubtest homogeneity, it received half credit because there was no evidence of a pool of items or selection from a pool. The *RCBA* received half credit because selection from a pool was not specified.

On the whole, *tester qualifications* were not well specified, especially for the tests that are designed for use with aphasic persons. The *Auditory Comprehension Test for Sentences* (*ACTS*) received half credit because there is some reference in the manual to speech-language pathologists, and there is also a statement that the test "presents no difficulty to the examiner familiar with testing aphasic individuals" (Shewan, 1979, p. 21).

If a supplemental aphasia test met the *test interpretation* criterion, it meant that an interpretation of the scores was put forth by the test developers; no judgment was made about the validity of the proposed interpretation. Although DeRenzi and Ferrari (1978) did not provide interpretations for *Reporter's Test* scores, this test received half credit because a cutoff score for aphasic performance was provided.

Most of the tests received half credit for the *directions* criterion. Few provided instructions to the patient or the examiner regarding guessing when unsure of a response.

With regard to *validity,* only the construct validity of these supplemental aphasia tests was evaluated. Some of the tests that did not pass this criterion did provide information about content and/or concurrent validity in the test manual. The validity data for the *Coloured Progressive Matrices* (*CPM*) are incomplete; we did not have access to Part 3 of the latest version of the manual, which apparently does report reliability and validity information.

In terms of *score stability,* most of the tests received half credit or none at all. The two tests that passed this criterion fully were the *Peabody Picture Vocabulary Test–Revised* (*PPVT-R*) and the *NRST,* which provided test-retest reliability data for children up to the age of 18 or up to grade 9. Thus, these two tests do not provide score stability information for normal older adults or aphasic persons. The *RTT* received half credit because test-retest reliability data are provided for only five brain-damaged subjects. The *CPM* data are not complete for the reason mentioned above. The *ACTS* received half credit because there was no indication of age or time post-onset for the 15 aphasic persons for whom there is test-retest data. Although the *RCBA* manual reports some findings from two studies of its test-retest reliability, information about subject characteristics is scanty.

The final psychometric criterion assessed the adequacy of a *normative sample.* Again, the two tests that met this criterion fully (*PPVT-R* and *NRST*) were not normed on an older adult population, and there are no norms for aphasic adults. The *ACTS* had some of the better norms; the performance of persons with various aphasia types across a wide age

range was described. However, full information about the subsample of patients for which norms are provided was lacking. The *RTT* also provided norms for aphasic and non-brain-damaged subjects over a wide age range, but education and socioeconomic status (SES) were not reported. The *CPM* received half credit because the non-brain-damaged older adult norms did not include information about education, gender, or SES. The *BNT* received half credit because the non-brain-damaged adult norms (which are provided up to age 59) were described by age or education (not age and education), and there were no gender or SES data. The *BNT* aphasia norms did not include information about age, gender, education, or SES. For the *Word Fluency Test*, there was no description of age, gender, education, or SES for 30 non-brain-damaged controls or for a small group of nonaphasic left-hemisphere-damaged subjects. The *Reporter's Test* fell short on gender and SES information for a large group of aphasic and control subjects. In addition, the aphasic sample was not described adequately.

DISCUSSION

Ideally, information about a test's psychometric properties would be published in the test manual when the test becomes available for circulation. As this very often is not the case, a few courses of action are available: (a) use the test anyway, (b) shelve the test, or (c) attempt to improve or gather more information about the test's psychometric properties.

Nicholas, Brookshire, MacLennan, Schumacher, and Porrazzo (1989) published an exemplary post hoc study, in which they reported revised administration and scoring procedures for the *BNT* as well as normative data for older non-brain-damaged adults.

On a practical note, we offer some suggestions for disseminating post hoc information about a test's psychometric properties. It is imperative that test users have access to complete information about the adequacy of the assessments they rely on. Ideally, test authors would update the manuals routinely. Failing that, they could provide the test publisher with addenda to be supplied with a test package on a regular basis. These might include a summary of the relevant available studies, a reference list so that test users could track down the information, and/or reprints of the articles themselves.

REFERENCES

American Psychological Association. (1974). *Standards for educational and psychological tests.* Washington, DC: American Psychological Association.

Borkowski, J. G., Benton, A. L., & Spreen, O. (1967). Word fluency and brain damage. *Neuropsychologia, 5,* 135–140.

Brookshire, R. H. (1986). *An introduction to aphasia.* Minneapolis, MN: BRK Publishers.

Conoley, J. C., & Kramer, J. J. (Eds.). (1989). *The tenth mental measurements yearbook.* Lincoln: University of Nebraska Press.

DeRenzi, E., & Ferrari, C. (1978). The Reporter's Test: A sensitive test to detect expressive disturbances in aphasia. *Cortex, 14,* 279–283.

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

Goodglass, H., & Kaplan, E. (1983). *The Boston Naming Test.* Philadelphia: Lea & Febiger.

Hanna, G., Schell, L. M., & Schreiner, R. (1977). *The Nelson Reading Skills Test.* Chicago: Riverside Publishing.

LaPointe, L. L., & Horner, J. (1979). *Reading Comprehension Battery for Aphasia.* Tigard, OR: C. C. Publications.

McNeil, M. R., & Prescott, T. E. (1978). *Revised Token Test.* Baltimore, MD: University Park Press.

Muma, J., Pierce, S., & Muma, D. (1983). Language training in speech-language pathology: Substantive domains. *Asha, 25,* 35–40.

Nicholas, L. E., Brookshire, R. H., MacLennan, D. L., Schumacher, J. G., & Porrazzo, S. A. (1989). Revised administration and scoring procedures for the Boston Naming Test and norms for non-brain-damaged adults. *Aphasiology, 3,* 569–580.

Raven, J. C. (1965). *The Coloured Progressive Matrices.* New York: Psychological Corporation.

Shewan, C. M. (1979). *The Auditory Comprehension Test for Sentences.* Chicago: Biolinguistics Clinical Institutes.

Skenes, L. L., & McCauly, R. J. (1985). Psychometric review of nine aphasia tests. *Journal of Communication Disorders, 18,* 461–474.

Wertz, R. T. (1985). Neuropathologies of speech and language: An introduction to patient management. In D. F. Johns (Ed.), *Clinical management of neurogenic communicative disorders* (pp. 1–96). Boston: Little, Brown.

APPENDIX 8.A
PSYCHOMETRIC CRITERIA USED IN TEST DEVELOPMENT

1. *Test Development.* "A test manual should describe fully the development of the test: the rationale, specifications followed in writing items or selecting observations, procedures, results of item analysis or other research" (p. 11).

2. *Tester Qualifications.* "The test manual should identify any special qualifications required to administer the test and to interpret it properly" (p. 15).

3. *Test Interpretation.* "The manual should state clearly what interpretations are intended for each subscore as well as for the total test" (p. 17).

4. *Directions.* "The directions should clearly state such critical matters as instructions on guessing, time limits, and procedures for marking answer sheets" (p. 18).

5. *Validity.* "A manual or research report should present the evidence of validity for each type of inference for which use of the test is recommended" (p. 31).

6. *Score Stability.* "The test manual should indicate to what extent test scores are stable, that is, how nearly constant the scores are likely to be if a parallel form of a test is administered after time has elapsed" (p. 54).

7. *Normative Sample.* "Norms presented in the test manual should refer to defined and clearly described populations. These populations should be the groups with whom the tester will ordinarily wish to compare the persons tested" (p. 20).

Note: From *Standards for Educational and Psychological Tests* by the American Psychological Association, 1974. Washington, DC: American Psychological Association.

APPENDIX 8.B
DECISION RULES FOR PSYCHOMETRIC CRITERIA

1. *Test Development*
   General description of specifications followed in item-writing process
   Evidence suggesting the existence of a pool of items
   Description of item analysis or other procedures used to select final test items from the pool

2. *Tester Qualifications*
   Description of tester's required educational background
   Description of tester's required experience with the test instrument

3. *Test Interpretation*
   Suggested interpretations of the overall test results
   Suggested interpretations for all scores or subscores yielded by the test

4. *Directions*
   Directions to test taker that he or she need not be certain of a response to answer
   Directions to tester about time limits for each task and an explanation of rationale behind omitting or including time limits on specific tasks
   Tester's procedures for marking answer sheets

5. *Construct Validity*
   Description of test developer's construct of language abilities
   Explanation of test's relationship to that construct

6. *Score Stability*
   Evidence of stability for a population described with regard to age and communication status

7. *Normative Sample*
   For tests considered norm referenced, definition of populations with respect to three of the following four characteristics: age, SES, education, and/or gender