THE IDENTIFICATION OF LANGUAGE IMPAIRMENT IN THE
SELECTION OF SPECIFICALLY LANGUAGE-IMPAIRED
SUBJECTS

REBECCA J. McCauley
Department of Communication Science and Disorders, University of Vermont, Burlington

M. J. Demetras
Division of Special Education and Rehabilitation, University of Arizona, Tucson

This review focused on the methods used to identify language impairment in specifically language-impaired subjects participating in 72 research studies that were described in four journals from 1983 to 1988. The single most frequent source of information used in the identification process was found to be test data. There was, however, considerable variability and, often, a lack of clarity regarding the specific number and identity of tests used. More specific findings on test use indicated that researchers routinely assessed both expressive and receptive language and that they used incomplete tests. When test scores used in identification and selection were examined, there was a wide range of score types, and age-equivalent scores were by far the most common and often the only type of score utilized. Conclusions are drawn regarding the impact of these findings on the interpretation and generalizability of this research literature.

KEY WORDS: specific language impairment, research methodology, subject selection, diagnosis

The valid identification of language impairment is a widely recognized and continuing challenge for subject selection in studies of language-impaired children (e.g., Lahey, 1988; Ludlow, 1980; Stark, Tallal, & Mellits, 1982; Tallal, 1987). Although this challenge confronts researchers studying any developmental disability that may include a language impairment (e.g., mental retardation, autism), it is particularly significant for researchers interested in specific language impairment (SLI). The identification of impairment in these children (and hence the selection process) is complicated by the heterogeneity of language skills such children present (e.g., Aram & Nation, 1975; Stark & Tallal, 1988). It is further complicated by the diversity of etiological perspectives guiding researchers who are interested in this population of children (e.g., Leonard, 1987) and are likely to define subjects differently according to their view of etiology. Still other factors complicating the selection process include the varied operational definitions used by researchers (Tallal, 1987) and the varying performance and adequacy of measures incorporated within those definitions (Demetras, 1984; Fuchs, Fuchs, Benowitz, & Barringer, 1987; McCauley & Swisher, 1984a). Continued progress in the understanding and treatment of SLI depends on continued improvement in subject selection and description.

Historically, diverse problems in subject selection and description have plagued research across a wide range of language disorders. A review of three methodological studies demonstrates that the procedures used to identify and describe the targeted language impairment represent a recurring and significant threat to valid subject selection. In the first of these three studies, Brookshire (1983) reviewed 52 articles concerning adult aphasic subjects, written over approximately 10 years and appearing in four journals. Although he primarily faulted the articles for consistently failing to describe subject characteristics known to affect language functioning (e.g., age, severity of aphasia), he also criticized the frequent failure to operationalize subject subgroupings (e.g., Wernicke’s aphasia, Broca’s aphasia, etc.). According to Brookshire, benefits from more consistent reporting of subject selection methods and more precise subject descriptions include helping to resolve apparent conflicts in the research literature and to clarify the generality of research findings to individual clients.

Similar concerns about subject selection and description were voiced in a review of research on autism (Kistner & Robbins, 1985). Kistner and Robbins examined 255 articles, most drawn from 1978 to 1982, and found that the majority (78%) reported some information about subject selection criteria, usually including citations to several standard definitions of autism (e.g., Rutter, 1978). However, they found that very few investigators provided information about methods used to operationalize standard definitions. In addition, they found that very few provided information about the reliability of the selection process and only 6% used quantitative measures in selection. Kistner and Robbins reported a trend in at least one journal toward greater clarity in subject selection. They argued, however, that still greater attention to subject selection and description is needed because of the heterogeneous nature of autistic subjects and because of growing interest in the delineation and study of subgroups.

In the paper most closely related to the present undertaking, Wickstrom, Goldstein, and Johnson (1985) examined subject descriptions of language-disordered children in 40 intervention studies conducted from 1978 to 1983 that appeared in seven different journals. They
looked at studies of SLI, autistic, mentally retarded, and other language-impaired persons under the age of 16. When 75% was taken to represent widespread or "standard" use, the authors found that only age, sex, and diagnosis were standard components of subject descriptions.

With regard to the identification of language impairment, Wickstrom et al. found that although 83% of the articles identified the diagnostic label of subjects, only 28% described the specific criteria used to select subjects. In addition, in 55% of the articles, language performance was described globally, without reference to performance within specific modalities or domains. This is disconcerting, given that the focus of all the studies was the efficacy of treatments undertaken to modify language performance. Wickstrom et al. ended their paper with recommendations for factors to be considered when describing subjects; language assessment and selection criteria were at the top of the list.

An exception to the vagueness in subject selection methodologies suggested by the Wickstrom et al. review is offered in an influential paper by Stark and Tallal (1981). As part of a project designed to examine sensory, perceptual, and other neuropsychological abilities in three groups of communication-disordered children (Stark & Tallal, 1988), Stark and Tallal developed an explicit methodology for selection of SLI children. In addition to a careful characterization of exclusionary criteria (e.g., presence of hearing impairment, neurologic disorders, etc.), Stark and Tallal offered a operational definition for children aged 4 12-8 years for language impairment of sufficient severity to pose a risk of significant learning disorder. This definition consisted of a test-derived Language Age for each child (based on the mean of Expressive and Receptive Language Ages) falling at least 12 months below either chronologic or Performance Mental Age, whichever was lower. In addition, the SLI children had to have a Receptive Language Age at least 6 months below Performance Mental Age and an Expressive Language Age at least 1 year below Performance Mental Age.

Seen with almost a decade of hindsight, Stark and Tallal's definition of language impairment has some methodological limitations. First, it is stated in terms of age-equivalent scores. The use of this type of score undermines the probable reliability of the definition even as it deceptively appears to permit comparisons between language modalities and between language and other areas of cognitive performance. This limitation was almost certainly necessitated by a scarcity of tests in 1981 that would permit use of a more adequate score type across all tests. A second limitation, clearly acknowledged by the authors, was the exclusion of children with certain patterns of impairment (e.g., children with significant expressive language deficits but normal or near-normal receptive language). Thus, the definition includes only certain patterns of language impairment in SLI children and not others. These limitations are described in order to demonstrate one of the principal values of clearly operationalized selection criteria—their accessibility to scrutiny and, therefore, to both straightforward interpretation and the possibility of systematic alteration.

In summary, problems in the selection and description of various language-disordered populations have often centered on the inadequacy or ambiguity of the methods used. In the literature reviews just cited, the methods used to identify the existence of a language impairment have been particularly problematic. The purposes of this paper are to summarize the assessment procedures used to identify a language impairment in research on SLI over a 6-year period, to investigate the extent to which tests are used in that process, and to offer suggestions about ways in which to improve identification and selection of SLI subjects.

**METHOD**

**Review Process**

Articles were examined from four journals: *Applied Psycholinguistics, Journal of Speech and Hearing Disorders, Journal of Speech and Hearing Research,* and *Language, Speech, and Hearing Services in Schools.* These journals were selected because they included at least six articles focusing on SLI subjects during a 6-year period. Articles were examined if they (a) appeared between 1983 and 1988, (b) reported on group or single subject research designs involving English-speaking subjects with SLI, and (c) did not focus on the identification process as such. This final criterion was intended to exclude articles in which a primary focus was the identification of language impairment (e.g., Ribner, Becker, Marks, Kahn, & Wolfson, 1983).

Variables used in the examination of each article addressed basic design information and the role of six sources of information in identification of a subject as language impaired: referral source, treatment history, direct observation by a speech-language pathologist, questionnaires, language sampling and analysis, and test data. These categories run the gamut from relatively subjective informal measures to more objective formal measures of language status. Together, they exhaust the probable sources of information used to identify a subject as language impaired.

During our rating of each article, we decided whether or not each of the six sources of information was used to identify language impairment in the article. The category referral sources was indicated when the researcher accepted the referral source's report of a previous diagnosis of language impairment as a major factor determining the child's group placement within the study. Although test data may have influenced diagnoses preceding such referrals, this category was used when there was no evidence of test use as a required precursor to the referral. A related category of information, treatment histories, was indicated when the researcher considered a potential SLI subject as language impaired on the basis of the clinical case history or school records indicating past or ongoing
treatment may have been an important part of a subject's case history or school records, this category was used rather than "test data" unless test data were identified as required for a subject to be considered language impaired. Direct observations were indicated when researchers included results of an observation period of the child's behaviors by a speech-language pathologist as part of the inclusion criteria. The category questionnaires indicated the use of standardized measures such as the Minnesota Child Development Inventory (Ireton & Thwing, 1974) as well as less formal, researcher-prepared questionnaires. For the category designated language sample data, specific information was recorded concerning the use of four specific procedures—standardized form, content, and use analyses (e.g., Bloom & Lahey, 1978); mean length of utterance (Brown, 1973; Miller & Chapman, 1981); type-token ratio (Brown, 1973); and the 14-morpheme count (Brown, 1973).

Test data represented the category of greatest interest for the purposes of this study. In addition to whether or not test data were used to identify language impairment, additional information was obtained about the nature of those data and how they were used in the decision process. Specifically, when test data were used, information was recorded concerning which specific tests were used and what types of scores were reported (e.g., raw scores, age-equivalent scores, z scores, etc.). Information regarding test use included whether the same type of score was used across tests, whether standardized test procedures were modified, and whether the tests used were appropriate for the age group. In addition, the test names were recorded to allow us to identify the language modalities tested and to examine the psychometric adequacy of tests most frequently used by researchers in the identification and selection process.

In addition to the six possible sources of information described above, a seventh variable, "other," allowed us to identify any unanticipated source(s) of information used in that selection decision. This category was never actually used because the first six variables describing sources of information proved exhaustive for the reviewed articles.

Frequently, it was difficult to determine whether subject characteristics cited in a method section were used to identify a child as language impaired for purposes of subject selection or, alternatively, whether such characteristics were used to describe language characteristics observed in subjects selected on other grounds. Although the distinction between selection criteria and descriptive variables may appear subtle, it is important because selection criteria are presumably those to which readers should most closely attend when considering the generalizability of study findings. Therefore, only those characteristics that were identified as having been used to make subject selection decisions were rated as "used" rather than "not used."

Seven articles, or approximately 10% of those reviewed, were independently rescored by the second author in order to obtain an estimate of interobserver point-by-point agreement (Kazdin, 1982) for ratings of "used" versus "not used" on each of the possible sources and specific testing practices examined. The second author's rating was obtained after a brief training session consisting of a discussion of judgment criteria and practice ratings on two articles similar to those included in this study. Interobserver agreement was 88% for all variables examined.

RESULTS

Seventy-two articles were considered eligible for review. Table 1 records the number of articles reviewed from each journal for each of the years studied. The majority of articles appeared in the Journal of Speech and Hearing Disorders and Journal of Speech and Hearing Research—44% and 31%, respectively. The remaining articles appeared in Language, Speech, and Hearing Services in Schools and Applied Psycholinguistics—15% and 10%, respectively. In all, 67 of the studies were experimental or descriptive group designs, and 5 were single subject experimental designs. One researcher and his colleagues were responsible for a total of 11 of the 72, or about 15% of the reviewed articles. Because these studies sometimes represented different analyses of the same children and it therefore seemed plausible that this body of work might be consistently unlike that of others, the presentation of results that follows will include mention of general findings that would be changed if those articles were excluded from consideration.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Applied Psycholinguistics</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Journal of Speech and Hearing Disorders</td>
<td>6</td>
<td>3</td>
<td>9</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>32</td>
</tr>
<tr>
<td>Journal of Speech and Hearing Research</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>Language, Speech, and Hearing Services in</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schools</td>
<td></td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>Year totals:</td>
<td>12</td>
<td>9</td>
<td>19</td>
<td>9</td>
<td>16</td>
<td>7</td>
<td>72</td>
</tr>
</tbody>
</table>

TABLE 1. Number of articles regarding specific language impairment published in four journals between 1983 and 1988.

Downloaded From: http://jsld.pubs.asha.org/ by a Ohio State University - Library User on 12/06/2015
Terms of Use: http://pubs.asha.org/ss/rights_and_permissions.aspx
Results are presented in four sections: (a) the sources of information used to identify language impairment, (b) frequency of test use, (c) the language modalities tapped in the identification process, and (d) the types of score(s) reported for the test data, when tests were used. These results exclude the description of methods used to identify language impairment in a non-English-speaking subject group (viz., a group of monolingual Italian speakers in Leonard, Sabbadini, Volterra, & Leonard, 1988) because the procedures used in the selection of this group would necessarily be quite different from those used for English speakers.

Sources of Information

Figure 1 summarizes the sources of information researchers used to determine language status. The most frequently used source was test data. Referral, language sample, and treatment history were used to identify language impairment less frequently, direct observation of possible subjects was used quite infrequently, and questionnaires were not used at all. None of the articles used information needing to be categorized as “other.”

The number of sources used by individual articles ranged from one to four different sources, with a mean and mode of two sources. Seventeen articles, or approximately 24% of those reviewed, used only one source of information to determine language status.

Frequency of Test Use

The frequency and identity of tests used in the identification process were examined because of the heterogeneity of language deficits observed in SLI children and the psychometric inadequacy of many measures used to assess these skills. Tests were used in 53 articles; the number of tests used in each article ranged from 0 to 6 or more. In 7 of the articles, researchers indicated that tests were used but did not identify the tests by name. In 10 of the articles, the precise number of tests used could not be determined because the researchers reported only partial lists of tests as examples of measures used in the identification process. Of the remaining 36 articles—those that permitted determination of the number of tests used—about half (47%) used only one test.

Table 2 summarizes the frequency with which individual tests were used. Thirty-one different tests were used in the 53 articles. The two most frequently used tests appeared in only 12 articles. In addition, 14 (45%) of the tests were used in no more than 1 article. Although much of this variability might have been due to the fact that different age ranges were used in different articles, an examination of the more frequently used tests suggests otherwise. Many of the tests were intended for test takers across a relatively wide range.

If we exclude the 11 articles contributed by one group of researchers, the frequency with which individual tests were used changed appreciably for only three tests. Specifically, the Preschool Language Scale (Zimmerman, Steinier, & Pond, 1979) was only used in 4 articles by other researchers and the Test for Auditory Comprehension of Language (Carrow, 1973), in only 9 articles. Similarly, the Illinois Test of Psycholinguistic Abilities was used in 8 rather than 10 articles, when these 11 articles were excluded. Other rankings, however, appeared only minimally affected by the practices of that group of researchers.

Some of the tests in Table 2 were not used in their entirety for the identification of subjects. In 14 of the 46 articles that named specific tests, researchers reportedly used individual or groups of subtests. Tests treated in this manner were the ITPA, McCarthy Scales of Children’s Abilities, NSST, PLS, TOAL, and TOLD. For two of these tests containing numerous subtests—the ITPA and the TOLD—the same four subtests appeared to be selected routinely: the Auditory Reception, Auditory Expression, Grammatic Closure, and Verbal Expression subtests of the ITPA, and the Picture Vocabulary, Grammatic Understanding, Oral Vocabulary, and Sentence Imitation subtests of the TOLD.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>PLS, TACL</td>
</tr>
<tr>
<td>11</td>
<td>PPVT</td>
</tr>
<tr>
<td>10</td>
<td>ITPA, NSST</td>
</tr>
<tr>
<td>9</td>
<td>PPVT—P</td>
</tr>
<tr>
<td>8</td>
<td>TOLD</td>
</tr>
<tr>
<td>7</td>
<td>CELI</td>
</tr>
<tr>
<td>4</td>
<td>CELF</td>
</tr>
<tr>
<td>3</td>
<td>ACLC, DSS, DTIA, TOAL</td>
</tr>
<tr>
<td>2</td>
<td>BLST, TACL—R, Token, TTC</td>
</tr>
<tr>
<td>1</td>
<td>RTBC, CELF—Screening, CLAM, DST, EOWPVT, HTLD, LSA, McCarthy, SICD, SPLIT, TOLD—I, TOLD—P, UTLD, Word</td>
</tr>
</tbody>
</table>

Figure 1. Sources of information used in 72 articles for the identification of language impairment.


Language Modalities

An examination of the tests that were identified by name provided the basis for determining the extent to which researchers examined both expressive and receptive language. This seemed especially important because 14 articles (30%) used only one test, thus almost requiring the use of a more comprehensive language test to avoid dependence on one modality in the identification of impairment. Forty articles (87%) examined both receptive and expressive language. In 6 articles, researchers examined only one language modality: 3 (6.5%) examined receptive language and 3 examined expressive language. As indicated previously, 26 articles either did not use tests in the identification process or did not name specific tests. In these cases there was no way of knowing which language modalities were assessed. Nonetheless, the overall frequency with which expressive and receptive language were explicitly described suggests higher standards of description than those suggested by Wickstrom et al.'s (1985) report that language performance was described without reference to language domains or modalities in 55% of the language treatment studies they examined.

Types of Scores

Types of scores used in the identification process were examined because of known differences in the psychometric adequacy of different types of scores (Anastasi, 1982; McCauley & Swisher, 1984b; Nitko, 1983). Table 3 summarizes the frequency with which different score types were reported for the 53 articles using test data. Some researchers reported only one type of score (e.g., age-equivalent scores), whereas others reported several (e.g., age-equivalent scores and percentile ranks). The word "unspecified" reflects cases in which researchers used tests but did not indicate the type of score used. In some cases researchers used more than one test but reported the type of score used for only one of them. In these cases, this pattern of scores is designated by the listing of one or more score types in combination with "unspecified."

Scores that include information about normative group variability (i.e., z scores and other standard scores) were used in 24% of the articles (11% individually and 13% in combination with other types of scores). In contrast, age-equivalent scores, which do not incorporate information about group variability, were the most frequently reported type of score. In 41.5% of the articles, age-equivalent scores were used alone; and in an additional 17%, they were used in combination with other types of scores. No other type of score—individually or in combination—approached the level of use of age-equivalent scores. In 15% of the articles, the type of score was unspecified.

If the 11 articles contributed by one research group are excluded, the percentages in Table 3 change for two categories, but without any effect on the relative frequency with which different score types were used. Specifically, when those 11 articles are excluded, age-equivalent scores only were used in 33% and age-equivalent scores plus percentile ranks in 7% of the remaining articles.

DISCUSSION

This discussion consists of a summary of major findings, a discussion of the implications these have for the generalizability and interpretability of SLI research, and a set of recommendations. This review focused on the methods used to identify language impairment in SLI subjects participating in 72 research studies described in four journals from 1983 to 1988. By far the single most frequent source of information used in the identification process was found to be test data. There was, however, considerable variability and, often, a lack of clarity regarding the specific number and identity of tests used. More specific findings on test use indicated that researchers routinely assessed both language modalities and that they used incomplete tests—that is, single or multiple subtests from a given instrument. This use of subtests was observed in approximately one-third of the 46 articles in which tests were named. When test scores used in the identification process were examined, a wide range of score types were used, but age-equivalent scores were by far the most common and were often the only type of score utilized.

Each of these findings has implications for the quality of the identification process used in the selection of subjects in recent research and thus on the quality of that literature as a whole. First, the frequent use of tests carries with it the potential for both positive and negative effects on the identification of language impairment. Potential positive effects of test use stem from their quantifiability and relative objectivity—features that might promote consistent implementation and thus increase the comparability of subjects across studies, which would allow for meaningful integration of research find-
ings. The potential negative effects of test use have been described at length (e.g., Kelly & Rice, 1986; Leonard, Prutting, Perozzi, & Berkley, 1978; Lieberman & Michael, 1986; McCauley, 1989; Newhoff & Leonard, 1983). The potential negative effect most relevant to this review is that there is very little evidence about the ability of such tests to differentiate normally developing children from language-impaired children—that is, very little evidence regarding concurrent validity (Howell, Skinner, Gray, & Broomfield, 1981; McCauley & Swisher, 1984a; Sommers, Erdige, & Peterson, 1978; Stark et al., 1982). Of those few studies that have attempted to provide such data, the results for individual and groups of tests are discrepant from one study to another (Demetras, 1990; Howell et al., 1981; McLoughlin & Gullo, 1984; Sommers et al., 1978). Thus tests used in the identification process may not live up to their promise of reliable and accurate identification of language impairment.

A more specific implication of the identification practices reviewed here concerns the administration of subtests in lieu of entire instruments. Although this is a common practice (and understandably so because of practical considerations), the norms assume questionable utility when the standard test procedures used to establish the norms are altered. Given this outcome, researchers must proceed cautiously and probably will want to avoid the practice of using subtests as a way to achieve more accurate subject identification and selection.

A second specific implication concerns the predominance of age-equivalent scores (language ages) in the identification of impairment of SLI subjects. Previously, in the context of a hypothetical clinical case, McCauley and Swisher (1984b) suggested that a common misuse of norm-referenced tests was overreliance on age-equivalent scores. Such developmental scores fail to take individual differences into account, tend to be less reliable than other score types, and are prone to misinterpretation (e.g., Anastasi, 1982; Petersen, Kolen, & Hoover, 1989). This predicted misuse has thus been documented by the current review of published research.

In addition to the possibility that known factors described thus far may negatively influence the quality of subject selection, numerous unknowns concerned us as we conducted this review. It was very difficult in many cases to determine whether specific procedures were used to select the children as language impaired or were merely included to describe their language skills. This lack of clarity hampers the replication of these studies. Furthermore, because of the wide variety of procedures used, the subjects in different studies could represent significantly different categories of children. Just as importantly, missing information regarding subjects prevents readers from generalizing these findings to specific clients.

To take the implications described above to their logical conclusions, the following recommendations for subject selection are made. First, in their descriptions of subject selection, researchers should explicitly differentiate subject selection from description—that is, they should distinguish between those subject attributes that determined eligibility for participation in the study from those that merely describe additional subject characteristics. Adherence to this recommendation should greatly diminish reader uncertainty regarding populations to whom findings can be generalized.

Second, given current limitations in assessment procedures, a “best practice” approach to identification of language impairment necessitates the use of more than a single source of information, as was done by most of the studies reviewed here. These limitations include the frequently discussed shortcomings of currently available tests, the lack of norms for most language-sampling procedures, and the wide variety of eligibility criteria used to place children in speech-language treatment. Clear description of the procedures used to integrate relevant information from the various sources used as part of the description of research methodology can provide a cohesive portrait of what constitutes language impairment among the children studied.

Third, when norm-referenced tests are used for the identification of impairment, researchers should follow scrupulously recommendations given for test administration and interpretation (e.g., American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1985; Anastasi, 1982). In particular, this recommendation requires researchers to consider the appropriateness of norm-referenced comparisons in light of subject characteristics and of the evidence available concerning the reliability and validity of the instrument for the identification of impairment in that population (Messick, 1989). Where relevant data are missing or unpromising, researchers should acknowledge the possible limitations associated with the use of the instrument or group of instruments for the purpose of identifying impairment—just as test users making clinical decisions have been advised to do (e.g., Anastasi, 1982).

As a fourth, more specific recommendation arising from this review, SLI researchers should consider the benefits of only using tests that allow them to operationalize their definitions of language impairment with cutoff scores described in terms of standard scores rather than age-equivalent scores. Use of standard scores facilitates comparisons of performance across language modalities and domains, as well as between language and other skill areas. Such comparisons are necessary, for example, for the examination of ability-achievement discrepancies (e.g., Berk, 1984).

At least one objection to this recommendation suggests itself. Specifically, the recommendation may seem to limit researchers to one or two language tests for certain age groups, thus equating the construct of language impairment to impairment on whatever is tapped in the content of that test or group of tests. In the study of learning disabilities, there is currently concern that the construct of intelligence, rather than being richly defined in the literature, has been impoverished by a nearly universal reliance upon a single measure, the Wechsler Intelligence Scale for Children—Revised (Wechsler, 1974) (e.g., Siegel, 1989). In the study of specific language
impairment, the restriction of test choice among researchers based on score type (as well as other issues related to validity and reliability, etc.) may similarly increase the prominence of some instruments. However, this prominence is likely to be accompanied by increased study of such instruments, by an increased pace of development of superior alternatives by members of the professional community, as well as by decreased use of instruments that fail to withstand even cursory examination—all of which are favorable alternatives to the current state of affairs.

Heterogeneity in specific language impairment imposed by variability in children’s language performances and competing etiological perspectives of researchers may never be resolved. However, the seemingly needless variability introduced by diverse selection procedures and ambiguous subject descriptions is directly under our control. Admittedly, subject selection methodology lacks the glamour of many other aspects of research design, and there are no accepted standards in this area. Nonetheless, without attempts to increase the clarity and adequacy of research practices in the study of specific language impairment, especially as these relate to the use of quantitative measures in subject selection, we risk drawing questionable conclusions about a poorly identified population.

ACKNOWLEDGMENTS

During the preparation of this manuscript, M. J. Demetras was partially supported by U.S. Department of Education Grant G00875001489.

The authors thank John Umbreit and John Brockmann for their critical comments on earlier versions of this manuscript and for their continued support. In addition, we would like to thank Laurence Leonard, Howard Goldstein, an anonymous reviewer, and Barry Guitar for their helpful comments and suggestions during the revision process.

REFERENCES


SIEGEL, L. (1989). IQ is irrelevant to the definition of learning.


Received August 25, 1989
Accepted November 7, 1989

Requests for reprints should be sent to Rebecca J. McCauley, Ph.D., Department of Communication Science and Disorders, University of Vermont, Allen House, Burlington, VT 05405.

**APPENDIX**

Thirty-One Measures Cited in Reviewed Articles


