

Samuel O. Ortiz  
*St. John's University*

## OVERVIEW

The concern with fair and equitable psychological assessment is a mirror of the values in our culture. As the nature of those values has changed, so have our concerns (Oakland & Laosa, 1976; Sandoval, 1998). The late 1950s and 1960s saw the rise of the civil rights movement, which brought to the forefront issues of discrimination that had remained entrenched but largely unspoken and unacknowledged in our nation's people and its systems. The rights of many diverse groups of people were upheld time and again by the U.S. Supreme Court, which affirmed equal protection under the law and the unconstitutionality of systematic and deliberate discrimination on the basis of race, creed, color, gender, or national origin (Irons & Guitton, 1993). During this dramatic shift in our country's cultural values, individuals with disabilities, including the parents of children with disabilities, were no less reticent to assert their rights (Jones, 1988; Laosa, 1976; Oakland & Laosa, 1976; Valdés & Figueroa, 1994). The passage of the landmark P.L. 94-142, the Education of Handicapped Children Act (later re-authorized and renamed the Individuals with Disabilities Education Act; IDEA) stands as testimony toward eliminating longstanding bias and discriminatory practices in education. IDEA contains a clear directive regarding the potentially discriminatory aspects of assessment, in that "tests and other evaluation materials used to assess a child...are selected and administered so as not to be discriminatory on a racial or cultural basis; and are provided and administered in the

child's native language or other mode of communication" (§300.532, Evaluation Procedures).

Attention to the attenuating effects of racial, cultural, and linguistic differences on test performance is evident in the wording. Moreover, this same section of IDEA requires the use of a "variety" of tools and strategies, emphasizing the fact that tests alone are unlikely to provide an accurate picture of functioning. Thus, nondiscriminatory assessment is not defined as a single procedure or test, but as a wide range of approaches that collectively seek to uncover as fairly as possible relevant information and data upon which decisions regarding functioning and performance can be equitably based. In other words, nondiscriminatory assessment is not a search for an unbiased test but rather a process that ensures every individual, not just those who are different in some way, is evaluated in the least discriminatory manner possible. To live up to its very name, nondiscriminatory assessment should be applicable to everyone, not just those from a particular group. In the most basic sense, "it is recognized that nondiscriminatory assessment may be considered one dimension of the more general problem of valid assessment of any child" (Oakland, 1976, p. 1).

Completely unbiased assessment, however, is an illusion. It is impossible to eliminate every single instance of bias or every potentially discriminatory aspect of assessment. Even with the most objective means and methods, bias exists in some form or another. Because elimination is unrealistic, the goal of nondiscriminatory assessment should be viewed as an effort to reduce it to the maximum extent possible.

Nondiscriminatory assessment is, however, much more than the sum of its parts. Although many discrete methods for reducing bias in assessment have been presented (Gonzalez, Brusca-Vega, & Yawkey, 1997; Gopaul-McNicol & Thomas-Presswood, 1998; Hamayan and Damico, 1991; Sandoval, Frisby, Geisinger, Scheuneman, & Grenier, 1998; Valdés & Figueroa, 1994) much less attention has been paid to the development of a broad, comprehensive framework for nondiscriminatory assessment. Application of a comprehensive framework designed to guide the general collection and interpretation of data in a systematic manner seems key to the process of equitable assessment. Engaging in assessment without the benefit of a guiding approach produces unrelated and disconnected activities and procedures that will likely fail to address important areas of bias and will produce data that are ambiguous and difficult to interpret. Nondiscriminatory assessment is much more than considering which standardized tools should be used and which should not. There is no simple answer or prescription, and standardized tests represent only one element of concern with bias. The use of one method or one procedure that may reduce bias is hardly enough to constitute unbiased assessment (Jones, 1988; Sandoval et al., 1998; Valdés & Figueroa, 1994). The use and application of a comprehensive, systematic framework comprising a broad range of methods and procedures is critical to engaging in best practices in nondiscriminatory assessment.

## BASIC CONSIDERATIONS

The increasing diversity of the U.S. population, especially in the public schools, has caused service and professional organizations to recognize the implications for change in both practice and training of its members. For example, in 1990, the American Psychological Association (APA) published *Guidelines for Providers of Psychological Services to Ethnic, Linguistic, and Culturally Diverse Populations* (American Psychological Association, 1990) in an effort to encourage psychologists to (a) consider the influence of language and culture on behavior when working with diverse groups, (b) consider the validity of the methods and procedures used to assess minority groups, and (c) make interpretations of resultant psychological data within the context of an

individual's linguistic and cultural characteristics (Lopez, 1997).

Many researchers have outlined certain procedures or methods related to compliance with these directives (e.g., Gonzalez et al., 1997; Gopaul-McNicol & Thomas-Presswood, 1998; Hamayan & Damico, 1991; Samuda, Kong, Cummins, Pascual-Leone, & Lewis, 1991; Sandoval et al., 1998; Valdés & Figueroa, 1994). Many of these methods revolve primarily around the modified use of or alternatives to standardized tests rather than specification of a comprehensive approach to nondiscriminatory assessment. The focus on reducing or avoiding bias related to the use of standardized, norm-referenced tests by these methods is not surprising given that tests are ubiquitous in psychoeducational assessment and often carry significant implications with respect to questions regarding diagnosis and intervention (e.g., special education eligibility). School psychologists are quite familiar and comfortable with this form of assessment, and it usually forms the bulk of their skills in assessment (Reschly & Grimes, 1995). However, the training of school psychologists often fails to provide sufficient competency regarding what might make their use biased or discriminatory and even less about how to use them in a less biased or discriminatory manner (Geisinger, 1998; see also Flanagan & Ortiz, in this volume).

Despite the need for information and guidance regarding the manner in which school psychologists may apply equitably or find alternatives to standardized tests, the varied and numerous recommendations offered in the literature for nondiscriminatory assessment have not gained wide acceptance in general school psychology practice. One possible reason is the lack of a comprehensive framework for nondiscriminatory assessment that guides the manner in which effective methods for reducing bias can be integrated into practice in a deliberate and systematic way. Indiscriminant and uninformed use of bias reduction techniques is a poor foundation upon which to claim fairness. Assessment activities that seek to reflect equity need to be undertaken within a broad framework built upon systematic and informed procedures that are brought together to form a valid context for interpretation. Apart from discussion of the main issues involved in nondiscriminatory assessment, the significance of the lack of a comprehensive and guiding model is underscored by

the delineation of “best practices” in the latter portion of this chapter that is presented within the structure of just such a framework.

### Hypothesis Testing

Nondiscriminatory assessment should incorporate the notion of hypothesis generation and testing. Although psychometric data are often viewed as objective, they have no inherent meaning and derive significance only from interpretation. Personal and professional bias often leads to idiosyncratic interpretations of the same data, in particular when assessment was begun with preconceived ideas. Consciously or unconsciously, bias on the part of the evaluator affects interpretive decisions and is known as confirmatory bias (Matsumoto, 1994; Sandoval, 1998). The chances of making incorrect inferences about data on the basis of preconceived ideas can be reduced through an approach that utilizes hypothesis generation and testing.

When a school psychologist or other evaluator conducts an assessment with pre-conceived notions regarding what the data will show (e.g., expected patterns of performance on a test), confirmatory bias can occur both in the type of data that are collected and the manner in which the data are interpreted (Matsumoto, 1994; Sandoval, 1998). For example, learning problems in the classroom may be erroneously ascribed to attention difficulties and thus subsequent data gathering efforts will tend to focus only on examining issues related to attention at the expense of other potential factors, such as limited English proficiency, that may be related to the learning problem. Conversely, attributions of behavior that are made on the basis of stereotyped or pre-conceived notions can often steer assessment away from the real cause of many school-related problems. Indeed, if there exists a belief that a student’s learning problems or behavior are attributable to personality, environmental, cultural, or linguistic differences, then no assessment may even be undertaken in cases where the learning difficulties may in fact be related to factors that can be readily ameliorated (e.g., instructional mismatch, health problems, sensory dysfunction, and so forth). These preconceptions can be particularly discriminatory whenever standardized tests are used. Believing that an individual is disabled can directly affect the manner in which the test is administered and scored

(benefit of the doubt will tend toward expectancies of dysfunction and disability) and the manner in which data are interpreted. The evaluator will tend to look for patterns and results in the data which support the pre-conception and is predisposed to perceiving only those patterns consistent with the a priori beliefs. Moreover, there is a tendency to ignore, minimize, or reject data counter to the assumption (Matsumoto, 1994; Sandoval, 1998).

Bias related to preconceived notions of dysfunction or discriminatory misattributions of performance or behavior tend to influence the very nature and range of data that will be collected. The questions asked during interviews, the behaviors observed in the classroom, and the work samples chosen for analysis will all be influenced by such preconceptions. Asking a parent when their child first learned to walk is different than asking if their child was in fact a late walker. School psychologists may reduce this form of bias by avoiding attempts to confirm presumptions of pre-existing deficits and testing hypotheses instead. The process of assessment should begin with the hypothesis that the examinee’s difficulties are not intrinsic in nature, but rather that they are more likely attributable to external or environmental problems. When standardized tests are used, the same assumption of normality should be used. In other words, the individual being tested is not impaired and general ability, performance, or functioning in any specific area is within normal limits. This assumption forms, in essence, a null hypothesis that can be evaluated with both quantitative and qualitative data to determine if it should be retained or rejected in favor of an alternative hypothesis (i.e., that performance is not average or within normal limits). When the process of evaluation is initiated with a presumption of normality, it reduces the tendency to search for data or “see” patterns of dysfunction where none may exist.

Another nondiscriminatory benefit of testing hypotheses that are not based on preconceptions is achieved by using the process for intervention in a proactive rather than reactive manner. For example, evaluation that seeks to determine the particular conditions under which a student’s learning may be improved or accelerated leads to the collection of data that are directly tied to intervention. The very purpose of the evaluation in such cases is to enhance learning rather than simply diagnose the causes of poor performance. Even when there may be a diag-

nostic component to evaluation, assessment should always be linked to intervention and the potential discriminatory influence of confirmatory bias can be reduced significantly when the focus is on identifying ways to improve school performance and learning rather than attempting to simply pinpoint the underlying cause of observed problems.

In sum, although it may be difficult not to entertain preconceived notions regarding the reasons for learning difficulties, particularly if efforts at intervention and treatment have failed, evaluation of data must remain squarely focused on whether hypotheses, not opinions, attitudes, or beliefs are or are not supported. Unless and until the data suggest strongly to the contrary, the null hypothesis that an individual's school problems are related to situational, not intrinsic variables, or that behavior or performance are normal and intact, must not be rejected. In addition, the alternative hypothesis should not be considered as providing *de facto* support for any preconception. The reasons why an individual may be having true problems learning in the classroom or why performance on any given test might fall outside of the normal range are numerous. Nondiscriminatory assessment seeks to ensure that the vast array of potential causes for learning difficulties, behavior problems, or low performance (e.g., low motivation, physical illness, anxiety, cultural or linguistic difference) have been ruled out as primary causes for any observed learning problems or patterns of deficit in the collected data. The act of developing and using hypotheses that affirm normality to guide the collection and interpretation of data remains central to reducing confirmatory bias and establishing defensible practices in nondiscriminatory (i.e., less discriminatory) assessment.

### **Cultural and Linguistic Competency**

Nondiscriminatory assessment represents a collection of approaches; each designed to systematically reduce bias within the broader framework. Cultural and linguistic competence is fundamental to that process. Cultural competence reflects a knowledge base of, or direct experience with, the values, attitudes, beliefs, and customs of a particular culture that can be used as both guide and context for collecting and evaluating any and all assessment data (Leigh, 1998). Psychologists need not be raised natively in a

particular culture in order to derive such competence, but the necessary skills will not develop by reading a book or taking a trip (Frisby, 1998). Those not fortunate enough to receive direct experience and education toward development of cultural competence from their training programs will need to embark on a focused process that includes a variety of professional development activities (Geisinger & Carlson, 1998; Leigh, 1998). In some cases cultural advocates from the community can assist in providing consultation regarding the particular aspects of culture that may be relevant to the evaluation.

Linguistic competence is reflected in two distinct ways: the ability to communicate effectively in an individual's native language (eliminating the need for an interpreter) and possession of a knowledge base related to first and second language development and instructional methodology and pedagogy (Sandoval & Durán, 1998). Possession of the ability to communicate effectively in an individual's native language does not automatically imply competency in first and second language development, instructional methodology, and pedagogy. Research has demonstrated that both cultural (not race, but acculturation) and linguistic (proficiency) differences are significant factors that can influence an individual's performance on psychological, language, and achievement tests (Comas-Díaz & Grenier, 1998; Cummins, 1984; Frisby, 1998; Sandoval et al., 1998; Valdés & Figueroa, 1994). Ortiz and Flanagan (1998) note, "mere possession of the capacity to communicate in an individual's native language does not ensure appropriate, nondiscriminatory assessment of that individual. Traditional assessment practices and all their inherent biases can be quite easily replicated in any number of languages" (p. 426). The entire process of assessment is subject to bias whenever there is a failure to account for culturally based influences including conceptions of time, world views, patterns of acculturation, normative behaviors, beliefs, values, attitudes, and expectations (Frisby, 1998; Salvia & Ysseldyke, 1991).

In general, the combination of cultural and linguistic competence may be defined as possession of the following: (a) skill and competence in selecting and using culturally appropriate methods, procedures, and tools that are designed to reduce bias systematically in assessment; (b) knowledge of, and familiarity with, cultural factors relevant to the indi-

vidual being assessed and the ability to evaluate data within the context of that culture; (c) knowledge of language development, second language acquisition, models of bilingual or English as a Second Language education and their relationship to achievement and school based learning; and (d) the ability to communicate effectively and competently in the native language of the individual being evaluated (Cummins, 1984; Hakuta, 1986; Krashen, 1985; Leigh, 1998). Within these general definitions, it is important to recognize that culture-specific knowledge and linguistic ability are secondary to the knowledge bases involving nondiscriminatory assessment practices and multi-language development and instruction. Linguistic or cultural similarity between examiner and examinee does not guarantee that the examiner possesses the requisite knowledge bases. A psychologist who has acquired the knowledge and skills described above is, with the assistance of an interpreter or a cultural advocate or both, much better equipped to conduct assessments that are far less discriminatory than an individual who possesses none of these skills but “matches” the child in terms of language or culture.

### Using Standardized Tests

Because their training may not have provided direct education and supervision in nondiscriminatory assessment or cultural and linguistic competency, school psychologists may resort to utilizing procedures and tests that are not suitable or appropriate for measuring cognitive abilities or intellectual functioning in equitable ways (Flanagan & Halsell Miranda, 1995; Geisinger & Carlson, 1998; Lopez, 1997). Use of standardized tests within the context of nondiscriminatory assessment requires knowledge of (a) the adequacy of representation of each norm or comparison group, (b) the full range of abilities that are being measured and those that are not, and (c) the inherent linguistic demands and cultural loading of each test (Flanagan & Ortiz, 2001; McGrew, Flanagan & Ortiz, 2000; Ortiz, 2001; Valdés & Figueroa, 1994). Whenever tests are selected, administered, and interpreted in a manner that is not systematic or guided by research, decisions and conclusions based on resulting data may be invalid or largely indefensible (Sandoval, 1998). In general, use of well-constructed, technically sound, native language tests (for

individuals who are evaluated in a language other than, or in addition to, English), where available, is preferable to tests with limited, poor, or unknown technical properties even if available in the native language (Geisinger, 1998). Oakland and Laosa (1976) reinforce this notion by stressing that “test misuse generally occurs when examiners do not apply good judgment or do not adhere to well-established professional procedures...governing the proper selection and administration of tests” (p. 17). In order to reduce bias that may arise from the use of standardized tests, knowledge regarding the ways in which such bias might operate and to what extent is required (Frisby, 1998; Valdés & Figueroa, 1994)

### Bias in Testing

The issues surrounding the nature of bias in standardized (in particular intelligence) tests have been discussed at length in the literature (e.g., Figueroa, 1990; Kamphaus & Reynolds, 1987; Oakland, 1976; Sandoval et al., 1998; Sattler, 1992; Valdés & Figueroa, 1994). With respect to the nature of bias, as operationalized in these investigations, the results have been rather unanimous: the majority of major intelligence batteries are not psychometrically biased (Valdés & Figueroa, 1994). Tests have been examined for bias related to item content, factor structure, mean group IQ differences, and prediction, all with the same result—no bias found (Jensen, 1974, 1980; Sandoval, 1979; Sandoval, Zimmerman, & Woo-Sam, 1983; Valdés & Figueroa, 1994). This is not surprising, however, because decades of test development have succeeded in creating instruments that measure quite well what they purport to measure. In the case of native English-speaking individuals raised in mainstream U.S. culture, they measure intelligence, or different facets of cognitive ability, or they predict achievement as well if not better than anything else irrespective of race or ethnic origin (Niesser et al., 1996). However, in the case of individuals whose experiential backgrounds (not race, ethnicity, or even culture per se) differ from the mainstream, bias may well operate. Oakland and Matuszek (1976) provide an eloquent synopsis of such bias and their comments deserve to be quoted in full:

The acculturation patterns governing the development of many children from racial-

ethnic minority groups or from lower socioeconomic homes also may be sufficiently different to warrant our judgment that the test is inappropriate. We must avoid the notion that all minority or lower socioeconomic children are, by definition, significantly different from those in the standardization sample. This position is prejudicial and unwarranted. However, we must be sensitive to the fact that important differences exist with respect to child-rearing practices, expectations and aspirations, language experiences, an availability of and involvement in informal and formal learning experiences, and that these and other factors may result in acculturation patterns which are not directly comparable to those which are more typical in the United States. The decision as to whether a child's acculturation patterns are similar to those generally reflected in the test's standardization sample can be made individually and only after a thorough knowledge of the child's background (p. 28).

Although bias has long been equated with differences in race, ethnicity, and culture, identifiable inequity lies less in these factors than it does in unique patterns of experience that may include either varying levels of acculturation or English language proficiency or both. The structure and design of intelligence and cognitive ability tests and the construction of representative norm groups are based on the notions of equivalency in levels of acculturation for both the individuals on whom the test was standardized and on whom the test will be used. In the assessment of any individual in today's diverse society, the validity of this assumption must be carefully evaluated. According to Salvia and Ysseldyke (1991), a fundamental principle within test development relevant to notions of bias is called "assumption of comparability." They write:

When we test students using a standardized device and compare them to a set of norms to get an index of their relative standing, we assume that the students we test are similar to those on whom the test was standardized; that is, we assume their acculturation is comparable, but not necessarily identical, to that

of the students who made up the normative sample for the test. When a child's general background experiences differ from those of the children on whom a test was standardized, then the use of the norms of that test as an index for evaluating that child's current performance or for predicting future performances may be inappropriate (p. 18).

The biasing effect from the use of psychometric instruments, therefore, operates whenever tests of intelligence and cognitive ability (developed and normed in the United States) are given to individuals whose cultural background, experiences, and exposure are not similar to or consistent with that of the individuals comprising the norm group against whom performance will be compared. Tests will likely measure a lower range of ability in such diverse individuals because they sample only the cultural content related to mainstream experience and not the full or range of cultural content possessed by the individual, and incorrect inferences may be drawn (Valdés & Figueroa, 1994). Tests may not be psychometrically biased, but they are culturally loaded and linguistically demanding to varying degrees (Sattler, 1992). In citing Jensen (1980), Frisby (1998) comments that "it is more accurate to characterize tests as falling along a continuum from 'culture reduced' to 'culture specific' or 'culture loaded.'" The same may be said for experience with respect to language proficiency requirements. Tests can be placed along a continuum from language reduced (i.e., nonverbal) to language embedded (e.g., a test of oral vocabulary). At the very center of nondiscriminatory assessment lies the need to recognize that incorrect inferences may be made on the basis of test scores that reflect performance as a function of measured variables not related to actual ability or aptitude. Moreover, reduction of bias in the use of tests can come only from knowledge regarding where a given test lies along these continua.

### **Native Language Testing**

Individuals involved in the assessment of linguistically diverse individuals should appreciate and recognize the difference between "bilingual assessment" and "assessment of bilinguals." The difference is not semantic, but rather represents two different research

traditions that have concomitant differences in application for nondiscriminatory assessment.

Bilingual assessment implies an approach to assessment that is conducted in a bilingual manner whereas assessment of bilinguals does not necessarily seek to gather information in a bilingual manner. Tests like the Bilingual Verbal Abilities Test (BVAT; Muñoz-Sandoval, Cummins, Alvarado, & Ruef, 1998) should be recognized as real, technical advances in “bilingual assessment.” Because this is an area of research that is relatively new, the underlying methods are neither complete nor without flaw. For example, despite the use of the term “bilingual” in its name, proficiency is still measured by the BVAT one language at a time which is quite different than the manner in which bilinguals are able to use both languages in an integrated manner (Bialystok, 1991; Grosjean, 1989). Aggregation of an individual’s language abilities into a bilingual composite after being measured separately is unlikely to be the most accurate operationalization of what bilingual ability actually is, yet it does manage to surpass previous methods in this respect. True bilingual sampling of an individual with standardized tests remains to be accomplished. Nevertheless, in time and with sufficient empirical support, these approaches to assessment may prove to have greater accuracy in measurement or practical utility. Indeed, they stand at the forefront of a relatively new line of research.

Assessment of bilinguals is the line of inquiry where the vast majority of research and practice exists and it has been conducted almost exclusively with tests given in English to people with varying levels of English language proficiency but rarely in a systematic way (Cummins, 1984). Therefore, a great deal more is known about how people who are non-native English speakers will perform on standardized tests given in English, than is known about how they perform on tests given bilingually or in their native language. The use of translated tests with unknown technical characteristics or limited norm group samples (e.g., WISC-R-M, WISC-R-PR) has provided little insight into the performance of bilinguals. In the future, the emergence of outstanding native language tests with superior technical qualities (e.g., Bateria-R, Woodcock, & Muñoz-Sandoval, 1996) will serve as the vehicles by which this distinct research question is addressed.

The distinction between bilingual assessment and assessment of bilingual individuals carries important

and distinct implications for nondiscriminatory practices. To engage in bilingual assessment, a psychologist must possess the requisite linguistic (and cultural) competency. Even when the competency requirement is met, the potential for bias is not fully diminished because there are no established procedures or guidelines to guide the process in a fair and equitable manner. How children growing up bilingual and bicultural in the United States should perform on tests that are normed on children raised in monolingual, single culture environments is unknown. Consequently, the vast majority of assessments to be conducted by psychologists will fall under the tradition of assessment of bilinguals, where the cultural and linguistic knowledge bases can be applied systematically within the context of a comprehensive framework for nondiscriminatory assessment.

## **BEST PRACTICES IN NONDISCRIMINATORY ASSESSMENT**

Apart from the considerations that have been discussed, in order to draw valid and defensible inferences from assessment data, nondiscriminatory assessment practices should be multi-faceted and guided by a comprehensive framework that integrates efforts to reduce bias in a cohesive and systematic manner. For any such framework to prove useful it must be practical and easily accommodated within the school or other applied setting. The following framework is designed to meet these criteria and is offered as a reasonable and efficient means for accomplishing reduction in the various aspects of the assessment process and its various steps are summarized in Table 1. The framework represents an initial attempt to coalesce the more salient and promising procedures and recommendations for nondiscriminatory assessment offered by both researchers and practitioners in school psychology and related fields. The framework is both linear and recursive in that a return to already completed steps in the process may well be necessary as new data are uncovered and new hypotheses formed, evaluated, and re-evaluated. The framework accommodates both individual and assessment team activities. Collaborative assessment, where members of an assessment team (including parents) work together and where information is shared and decisions rendered jointly, significantly improves the likelihood of success of any and all nondiscriminatory efforts.

**Table 1. A comprehensive framework for nondiscriminatory assessment**

1. Assess and evaluate the learning ecology.
2. Assess and evaluate language proficiency.
3. Assess and evaluate opportunity for learning.
4. Assess and evaluate educationally relevant cultural and linguistic factors.
5. Evaluate, revise, and re-test hypotheses.
6. Determine the need for and language(s) of assessment.
7. Reduce bias in traditional testing practices.
8. Utilize authentic and alternative assessment procedures.
9. Evaluate and interpret all data within the context of the learning ecology.
10. Link assessment to intervention.

**1. Assess and evaluate the learning ecology.** School psychologists should recognize that there exist an infinite number of reasons why any given individual is having learning difficulties and that intrinsic factors form only a small fraction of these possibilities. Nondiscriminatory evaluation begins with directing initial assessment efforts toward exploration of the extrinsic causes that might be related to any observed learning difficulties. Hypotheses should be developed that revolve around the individual's unique experiential background within the context of the learning environment. When assessment is conducted on culturally and linguistically diverse individuals, in particular, there are many reasons related to these experiential factors that can adversely affect classroom performance or behavior.

Very often it is the systemic interaction between these factors and those that exist in the learning environment that simply do not or are not able to accommodate them that creates a mismatch between instruction delivered and instruction needed. Although cultural or linguistic differences are probably two of the most common factors that are evaluated relative to the learning ecology, they are by no means the only ones and the learning ecology should not be ignored simply because a student's background is not characterized by diversity on these two dimensions. In addition, a student's learning ecology should not be thought of as being restricted solely to the classroom environment. Although focus on the classroom environment is central to the evaluation of learning problems, students learn a great many things in contexts other than the classroom. Comparison of behavior, performance, or functioning between these contexts (e.g., physical education, non-academic instruction, recess, home, community) is crucial in conducting nondiscriminatory assessment.

Data that inform evaluation of hypotheses related to ecological and systemic factors may be obtained via a variety of methods, including review of educational records, direct observation of instruction and teaching, review of the content, level, relevancy and appropriateness of the curriculum, analysis of the match between the curriculum and the student's needs, interview with parents, teachers, or the individual, and medical records. Often culture and its concomitant experiences most dictate the unique history of an individual but they should not be equated to or measured simply by skin color or ethnic heritage. Evaluation of cultural difference should be viewed as examination of relatively unique circumstances or learning experiences that can not be considered comparable to the experiences of individuals raised in the U.S. mainstream. Some examples that may be consistent with this definition include poverty, deafness or other disability, bicultural students, students with childhood trauma or abuse, and students from dysfunctional families. Although culture tends to be the major factor that influences an individual's development, it need not be thought of as being neatly circumscribed. Evaluation of the extent to which a student's experiences differ from that of mainstream students may not even be a function of culture, but simply the result of unusual or highly idiosyncratic experience. Specification of hypotheses should be "null based" (i.e., performance, behavior, or learning problems are due to extrinsic factors such as differences in experience, not intrinsic factors like ability) in order to prevent bias in the collection and interpretation of data.

**2. Assess and evaluate language proficiency.** In cases where the individual is a dual-language learner, it will be necessary to determine current levels of language



proficiency in both languages, especially with respect to Basic Interpersonal Communicative Skills (BICS) and Cognitive Academic Language Proficiency (CALP) (see Cummins, 1984). This information is often necessary in order to properly evaluate many of the ecological elements of the learning context related to dual-language experiences. Questions regarding opportunity to learn, the level and manner of instruction, the curriculum's linguistic relevancy and appropriateness, expected level of functioning or performance relative to English language development, and others are answerable only with this information. Thus, knowledge of an individual's proficiency in each language is crucial to nondiscriminatory assessment and interpretation. It provides the required context within which academic difficulties can be properly evaluated and forms the basis for the development of instructional interventions that are appropriate linguistically. Such data are usually gathered through any one of the various, standardized language tests available on the market today. Of particular note is the BVAT (Muñoz-Sandoval et al., 1998), which is available in 15 different language versions. The data may already exist in school districts where there are bilingual or English as a Second Language (ESL) programs. If the information is available but is out of date (older than 6 months), then new data should be collected. The impact of language experiences is profound, particularly in the preschool and early primary grades, and exposure to a second language no matter how small can have a significant impact on patterns of academic performance and basic skills development or acquisition (Cummins, 1984; Krashen, 1985). As such the presence of a language other than English in the home should not be minimized in the face of even considerable exposure to the English language.

**3. Assess and evaluate opportunity for learning.** The school setting provides perhaps the most significant context for learning. However, it is by no means perfect, and an individual may become a "casualty" of the educational system's failure to provide an effective or appropriate instructional program. The educational institution itself, including the curriculum, personnel, policies, and even the instructional setting, must be carefully evaluated to determine whether the individual has been provided with adequate "opportunity to learn" particularly in the case where significant cultural or linguistic differences exist. Again, the usual methods for collecting this type of data

include evaluation of classroom environment and teaching methods, direct observation of academic performance, review of the content, level, relevancy and appropriateness of the curriculum, analysis of the match between the curriculum and the student's needs, interviews with parents, current and previous teachers, interview with the individual, and review of existing educational records and progress reports. Specific factors that should be examined include: regularity of school attendance; experience with the school environment and setting; match between individual's native language and the language of instruction; parent's ability to support language of instruction; years (duration) of instruction in the native language and English; quality of native language and English instruction in ESL or bilingual programs; cultural relevance of the curriculum; frequency of changes in schools; relative consistency in and across curricula; teaching strategies, styles, attitudes, expectations; system attitude regarding dual language learners; and socialization with peers versus isolation from peers.

**4. Assess and evaluate educationally relevant cultural and linguistic factors.** Learning does not take place only in school but occurs throughout the broad scope of an individual's complete social milieu. Many factors outside and apart from the educational setting can significantly affect the learning process, and careful evaluation of the extent to which such factors might be present and might be affecting learning in the school is necessary in order to evaluate data from a nondiscriminatory standpoint. In cases where the individual is culturally or linguistically diverse, it will be necessary to assess and evaluate the experiential aspects of these particular variables and their relative influence on school-based learning, language development, and educational progress. The effect of small amounts of exposure to two or more cultures or languages during early childhood development may create circumstances that cause the individual to have experiences that differ markedly from that of other individuals within the U.S. mainstream and that can negatively affect school performance. In short, the ability to draw valid conclusions regarding school dysfunction from the whole of assessment data rests squarely on proper identification and understanding of the individual's total linguistic history as well as other factors that may have influenced the develop-

ment of both languages. This information is most commonly collected via observations across multiple settings, interviews with parents, teachers, and the individual, and review of existing educational records. Additionally, home visits are particularly effective for gathering this type of data. Factors that should specifically be examined include: current language(s) of the home, the individual's initial or primary language, the individual's total informal experience with the native language and English, individual's birth order or relative impact of siblings and their language development, individual's fluency in the native language and English, individual's and parent's level of acculturation, parent's fluency in the native language and English, parent's level of literacy in the native language and English, parent's level of education, and parent's socio-economic status.

**5. Evaluate, revise, and re-test hypotheses.** School psychologists should ensure that all reasonable and viable factors that could be related to an individual's observed learning difficulties have been thoroughly evaluated and ruled out as the "primary" cause of them. Within the school setting, it is only when enough confidence exists in the belief that there are no plausible or demonstrable external factors that can account for an individual's learning difficulties and that consideration of possible intrinsic factors should be entertained. This is necessary primarily to reduce confirmatory bias. Additionally, there should also be evidence that systematic and appropriate efforts to improve the student's classroom performance were undertaken but proved to be unsuccessful. The litmus test here is whether there is one or a combination of external factors present that can be reasonably presumed to be the primary cause of the individual's learning difficulties. If so, then the null hypothesis regarding normal behavior, performance, or average functioning (albeit, inhibited by the identified external constraints) should be maintained and the individual's learning problems should not be ascribed to intrinsic factors. The reduction of potentially discriminatory attributions regarding learning, behavior, or performance is further attenuated by returning to the development of additional interventions to address academic need within the classroom setting. In some cases, however, external factors may be present, but might only be contributing to and not directly causing the student's observed learning prob-

lems. When difficulties learning in the classroom setting or behavioral problems can not be reasonably ascribed to the primary influence of any such extrinsic factors, assessment may proceed appropriately to explore potential intrinsic factors with confidence that the process is operating in a fair and equitable manner.

**6. Determine the need for and language(s) of assessment.** The course of assessment may be significantly affected in cases where the individual is not a native English speaker (e.g., is limited English proficient) or uses an alternative mode of communication. This is particularly true for special education evaluations where IDEA mandates that assessors consider the individual's primary language ability (in addition to his or her ability in English) in the development of the assessment plan (§§300.532-330.534(b)). As assessment moves logically toward the use of standardized tests it becomes important to recognize that testing need not be conducted solely in the primary language or English. Moreover, nothing in IDEA or any other regulatory guideline mandates "parallel" testing in both languages. Exactly what should be assessed and in what language it will be assessed are decisions that rest with the assessor or assessment team. Factors that affect the selection of linguistically appropriate tools and techniques come from examination and review of existing pre-referral data, the unique background variables of the individual, and relevant referral concerns. Because each case is unique, and because the foundation of IDEA rests on the notion of "individualization" in both evaluation and instruction, it is inappropriate to make specific guidelines or rules with respect to decisions about the most appropriate language or combination of languages for testing. Within the framework of nondiscriminatory assessment, these decisions may be guided by the following general statements that represent only the most basic of guidelines: (a) individuals who are not proficient in English should be assessed in their primary language or native mode of communication in addition to any English language testing that may be appropriate; (b) individuals who are proficient in English may be assessed in their primary language or native mode of communication in addition to any English language testing that may be appropriate; and (c) all individuals, whether proficient in English or not, whose histories and backgrounds are not comparable to the U.S.

mainstream, should be evaluated by an assessor who possesses knowledge regarding the factors relevant to the individual's unique experiences and how they may affect learning and development.

A criticism often leveled at nondiscriminatory assessment suggests that it is not practical and involves considerable expenditures in terms of time and effort. To the contrary, the six steps that have already been delineated above can be accomplished well within the scope of any pre-referral intervention and assistance process. Use of a pre-referral process coupled with attendance at such meetings by professionals with the relevant knowledge and competencies (school psychologists are very well suited for this role) creates a system that is highly efficient by eliminating time wasted conducting inappropriate evaluations. Not only does creation and involvement in an effective pre-referral process reduce time and effort spent in evaluations, but it also helps to streamline compliance with legal mandates by facilitating scheduling and documentation. Nondiscriminatory assessment need not wait to begin upon formal referral.

**7. Reduce bias in traditional testing practices.** As discussed previously, bias in traditional testing occurs primarily when individuals whose backgrounds and experiences differ significantly from those on whom the test was normed. This is not an uncommon situation, and the issues of acculturation and English language proficiency significantly affect the validity of interpretations drawn from results of performance on such tests. Even where native language tests are available, potential bias remains. The process of nondiscriminatory assessment in using tests is represented by two distinct options: (a) administer tests in a standardized way and attempt to evaluate the results in a nondiscriminatory manner or (b) modify the testing process in a way that is less discriminatory initially.

In pursuing the first option, maintaining standardization has the advantage of allowing application of systematic methods for reducing bias. This includes use of existing (Mercer, 1979) or locally developed pluralistic norms that provide more appropriate comparison groups or use of information regarding the linguistic demand and cultural loading dimensions of the tests given (Flanagan, McGrew, & Ortiz, 2000; Flanagan & Ortiz, 2001; Ortiz & Flanagan, 1998). Because there is considerable research surrounding the nature and manner in which various groups of

diverse individuals perform on a wide range of available English language tests, maintaining standardization provides a foundation for nondiscriminatory assessment based directly on this body of empirical evidence. The CHC Culture-Language Matrix developed by Flanagan and Ortiz (2001) capitalizes directly on this research and provides practitioners with a systematic method that can assist in determining the relative impact of cultural and linguistic differences on test performance. Deviations from standardization would produce results that are unknown and unpredictable, as does use of native language tests for which an extensive literature base does not exist. Both approaches within this option provide the means for basing interpretation relative to more representative peers as well as expected patterns of performance as a function of acculturation and language proficiency. Moreover, knowledge of test properties relative to cultural loading and linguistic demand creates a basis for test selection that may also be less discriminatory without violating standardization.

The second option in using standardized tests involves modification and adaptation in ways that attempt to reduce acculturative or linguistic bias directly. Any such modifications, however, represent significant violations of standardization and automatically impugn the validity and interpretability of obtained results. The major drawback in applying modifications is the elimination of empirically established baselines for comparison or performance. Unlike performance on tests administered in English following standardization, how much any given modification or adaptation affects performance is not well known or defined. Consequently, the major benefit that may be derived in using tests in non-standardized may rest more in the collection of qualitative versus quantitative data. Even in this respect administration and scoring should remain systematic.

Standardized norm-referenced tests, both in English and the native language, can be modified in a variety of ways including bilingual administration, use of extended and expanded instructions on sample items, mediation of concepts to ensure comprehension prior to administration of items, repetition of items to facilitate comprehension, extension or elimination of time limits, acceptance of alternative responses (e.g., in a different language, culture-specific responses, through non-verbal gestures), and

additional probing and querying of incorrect responses. When carefully adapted and interpreted within the context of the individual's unique experiential background, the use of standardized tests can provide valuable qualitative information about functioning. Such data are often much more useful in instructional planning than the quantitative results that are derived. Despite the limitations, use of tests as tools that provide rich clinical or behavioral information should not be discounted as a basis for drawing less discriminatory inferences about ability or performance.

**8. Utilize authentic and alternative assessment procedures.** Nondiscriminatory assessment represents a collection of related and systematic procedures designed to reduce bias and implies the collection of a broad range of information. Non-standardized methods and information should not be excluded from the process. Whereas standardized, norm-referenced tests are driven mainly by questions and needs related to classification, diagnosis, and legal eligibility, authentic assessment is geared more toward answering questions regarding instructional needs and interventions—something that standardized tests do not address well. Examples of such procedures and methods include informal analysis of work samples, curriculum or criterion based measurement, performance-based assessment, portfolio assessment, and various test-teach-test frameworks such as dynamic assessment (Fischer & King, 1995).

When properly applied, non-standardized, alternative assessment strategies can provide valuable information especially in school-based and special education evaluations. In educational settings, authentic assessment often utilizes material that has already been provided through direct instruction. Evaluation of learning and performance through use of the curriculum-based materials and content reflects an authentic nondiscriminatory approach. Accordingly, comprehensive assessment should include information and data obtained through such methods. Examples include curriculum-based assessment or authentic measures of academic achievement and skill development, performance-based assessment that evaluates more by task completion within context than answering of factual questions out of context, criterion-referenced assessment using minimum levels or standards of performance, portfolio

assessment that documents development of skills learning and academic progress, informal analysis of actual work completed in the classroom setting, symbolic dynamic assessment of learning propensity using abstract stimuli, and authentic dynamic assessment of learning propensity using actual materials from the curriculum. Dynamic assessment is a particularly useful culture-reduced method that can be accomplished by using a wide variety of materials that provides relevant and useful information about performance that is directly applicable to instructional intervention and planning (Lidz, 1997).

**9. Evaluate and interpret all data within the context of the learning ecology.** All data collected in the course of nondiscriminatory assessment should be evaluated in an integrated manner utilizing the information obtained regarding the student's unique experience and background as the appropriate context. Knowledge of factors that may have played a part in creating significant differences between the experiences of the individual in terms of acculturation or language development provides the least discriminatory framework with which to evaluate and assign meaning to the patterns seen in the data. Although less obvious and more difficult to judge, information related to differences in acculturation are every bit as important as the more overt differences seen in language, and should not be minimized or ignored. Very often the meaning of the data will depend in large part on an understanding of the environmental influences (generally most associated with cultural and linguistic differences) that have transpired to shape the individual in unique ways and set the stage for observed and measured performance. In the final analysis, successful nondiscriminatory assessment is contingent upon application of this information. An effective method for ensuring equity in this process is based on the notion of convergence. The data collected in the course of assessment should cut across procedure or method and come together in a cohesive and convincing manner that supports the plausibility of final conclusions. Although in practice, a preponderance of evidence would be sufficient to provide validity to conclusions, care should be taken not to assign unwarranted importance or significance to any single piece of information or datum. Use of single scores, combinations or products of scores, and unduly favoring certain data over other data will lead

to discriminatory inferences and outcomes. Equivocal data support the null hypothesis that functioning is within normal limits and that any observed difficulties are the result of factors other than internal disability.

**10. Link assessment to intervention.** Assessment and evaluation are not interventions. Even when an individual has been assessed and it has been determined that no disability exists, the original problems that prompted the referral do not cease to exist automatically. Similarly, that a disability has been correctly identified does not make the best or most appropriate remedial methods instantly apparent. Assessment of any kind, nondiscriminatory included, is of little value unless it can be extended to incorporate appropriate interventions and treatment options. In school-based evaluations, modifications to the instructional program and the provision of specific remedial strategies are necessary whether the individual qualifies for special education services or not. Because the process of nondiscriminatory assessment has generated information regarding both relative performance as well as causal and contributory factors, it has considerable value in guiding the development of appropriate interventions and treatment strategies. As with use of the hypothesis-driven approach, awareness of the need to link assessment to intervention significantly affects the manner in which activities are conducted and the type of data that are collected. Failure to generate data for the purpose of intervention can be construed as the most discriminatory aspect of assessment as it affects all individuals equally.

## SUMMARY

No assessment is unbiased and no evaluation can ever be wholly nondiscriminatory. Attempts to be completely unbiased are doomed to failure. A more practical approach is to recognize the various sources of potential bias and use systematic procedures that will reduce it as much as possible. Nondiscriminatory assessment is not limited to individuals who are distinct from the cultural or linguistic mainstream. Rather, in keeping with the egalitarian implications of its name, nondiscriminatory assessment can provide fair and equitable evaluation of any individual regardless of background. Nondiscriminatory assessment is not a single tool or procedure that is applied

in isolation or without consideration of differences in individual experience and development. It is a collection of activities brought together in comprehensive fashion and used in a systematic way to address a variety of issues related to bias or discrimination. Nondiscriminatory assessment is broad and includes a wide variety of data generated from review of records, interviews, observations, standardized tests, and authentic methods. Bias is not a function of technical or psychometric deficiencies in tests but rather differences in experience between an individual taking the test and the individuals on whom the test was normed. Differences in race, ethnicity, or culture do not always reflect significant experiential differences. Direct training and education in nondiscriminatory assessment assists in drawing fair and correct inferences regarding patterns of, and reasons for, learning or behavior problems. There is no one right way to reduce bias, and there are no strict procedural specifications for which adherence is required. Nondiscriminatory assessment is, however, best carried out within the provisions of an overarching framework that brings bias reduction procedures together in a cohesive and logical manner and which assists not only in interpreting data fairly but also the collection of data in ways that are similarly less biased. Additionally, testing hypotheses regarding factors that are not child-centered but rather problem-centered and which affirm normality assists in reducing the discriminatory aspects of assessment and leads directly to the development of appropriate interventions. A general framework for nondiscriminatory assessment that achieves these goals was proposed. The process outlined in the previous section provides a structure whereby more than half of the components essential to nondiscriminatory assessment can be accomplished within the scope of general education and pre-referral activities. These activities provide data that inform the development of appropriate instructional interventions and guides subsequent assessment activities. Should assessment of learning or behavioral difficulties proceed to formal evaluation of the presence of an underlying disability, the remaining components in nondiscriminatory assessment that are often required in such cases can be completed efficiently in approximately the same amount of time and with no more effort than would be required in any other type of assessment. Ultimately, nondiscriminatory assessment should be “undertaken with

the intentions of improving children's development and helping persons make wise and informed decisions" (Oakland, 1976, p. 3). When data collection and interpretation are guided by responsive methods embedded in a systematic framework, the likelihood of fair and equitable decisions are increased.

## REFERENCES

- American Psychological Association. (1990). *Guidelines for providers of psychological services to ethnic, linguistic, and culturally diverse populations*. Washington, DC: Author.
- Bialystok, E. (1991). *Language processing in bilingual children*. New York: Cambridge University Press.
- Comas-Díaz L., & Grenier, J. R. (1998). Migration and acculturation. In J. Sandoval, C. L. Frisby, K. F. Geisinger, J. D. Scheuneman, & J. R. Grenier (Eds.), *Test interpretation and diversity: Achieving equity in assessment* (pp. 213–239). Washington, DC: American Psychological Association.
- Cummins, J. C. (1984). *Bilingual and special education: Issues in assessment and pedagogy*. Austin, TX: PRO-ED.
- Figueroa, R. A. (1990). Best practices in the assessment of bilingual children. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology II*. Washington, DC: National Association of School Psychologists.
- Fischer, C. F., & King, R. M. (1995). *Authentic assessment: A guide to implementation*. Thousand Oaks, CA: Corwin.
- Flanagan, D. P., & Halsell Miranda, A. (1995). Best Practices in working with culturally different families. In A. Thomas and J. Grimes (Eds.), *Best practices in school psychology III* (pp. 1049–1060). Washington DC: National Association of School Psychologists.
- Flanagan, D. P., McGrew, K. S., & Ortiz, S. O. (2000). *The Wechsler intelligence scales and Gf-Gc theory: A contemporary approach to interpretation*. Boston: Allyn & Bacon.
- Flanagan, D. P., & Ortiz, S. O. (2001). *Essentials of cross-battery assessment*. New York: John Wiley.
- Frisby, C. L. (1998). Culture and cultural differences. In J. Sandoval, C. L. Frisby, K. F. Geisinger, J. D. Scheuneman, & J. R. Grenier (Eds.), *Test interpretation and diversity: Achieving equity in assessment* (pp. 51–73). Washington, DC: American Psychological Association.
- Geisinger, K. F. (1998). Psychometric issues in test interpretation. In J. Sandoval, C. L. Frisby, K. F. Geisinger, J. D. Scheuneman, & J. R. Grenier (Eds.), *Test interpretation and diversity: Achieving equity in assessment* (pp. 17–30). Washington, DC: American Psychological Association.
- Geisinger, K. F., & Carlson, J. F. (1998). Training psychologists to assess members of a diverse society. In J. Sandoval, C. L. Frisby, K. F. Geisinger, J. D. Scheuneman, & J. R. Grenier (Eds.), *Test interpretation and diversity: Achieving equity in assessment* (pp. 375–386). Washington, DC: American Psychological Association.
- Gonzalez, V., Brusca-Vega, R., & Yawkey, T. (1997). *Assessment and instruction of culturally and linguistically diverse students with or at-risk of learning problems: From research to practice*. Needham Heights, MA: Allyn & Bacon.
- Gopaul-McNicol, S., & Thomas-Presswood, T. (1998). *Working with linguistically and culturally different children: Innovative clinical and educational approaches*. Needham Heights, MA: Allyn & Bacon.
- Grosjean, F. (1989). Neurolinguists beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36, 3–15.
- Hakuta, K. (1986). *Mirror of language: The debate on bilingualism*. New York: Basic Books.
- Hamayan, E. V., & Damico, J. S. (1991). *Limiting bias in the assessment of bilingual students*. Austin, TX: PRO-ED.
- Irons, P., & Guitton, S. (1993). *May it please the court*. New York: The New Press.
- Jensen, A. R. (1974). How biased are culture-loaded tests? *Genetic Psychology Monographs*, 90, 185–244.
- Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

- Jones, R. L. (Ed.). (1988). *Psychoeducational assessment of minority group children: A casebook*. Berkeley, CA: Cobb & Henry.
- Kamphaus, R., & Reynolds, C. (1987). *Clinical and research applications of the K-ABC*. Circle Pines, MN: American Guidance Services.
- Krashen, S. D. (1985). *Inquiries and insights: Second language teaching, immersion and bilingual education, literacy*. Englewood Cliffs, NJ: Alemany.
- Laosa, L. M. (1976). Historical antecedents and current issues in nondiscriminatory assessment of children's abilities. In T. Oakland (Ed.), *Non-biased assessment of minority group children: With bias toward none*. Paper presented at a national planning conference on nondiscriminatory assessment for handicapped children, Lexington, KY.
- Leigh, J. W. (1998). *Communicating for cultural competence*. Boston: Allyn & Bacon.
- Lidz, C. S. (1997). Dynamic assessment approaches. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 281–296). New York: Guilford.
- Lopez, E. C. (1997). The cognitive assessment of limited English proficient and bilingual children. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 506–516). New York: Guilford.
- Matsumoto, D. (1994). *Cultural influences on research methods and statistics*. Pacific Grove, CA: Brooks/Cole.
- Mercer, J. R. (1979). *System of Multicultural Pluralistic Assessment: Technical manual*. New York: The Psychological Corporation.
- Muñoz-Sandoval, A. F., Cummins, J., Alvarado, C. G., & Ruef, M. L. (1998). *The Bilingual Verbal Ability Test*. Chicago: Riverside.
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101.
- Oakland, T. (Ed.) (1976). *Non-biased assessment of minority group children: With bias toward none*. Paper presented at a national planning conference on nondiscriminatory assessment for handicapped children. Lexington, KY.
- Oakland, T., & Laosa, L. M. (1976). Professional, legislative, and judicial influences on psychoeducational assessment practices in schools. In T. Oakland (Ed.), *Non-biased assessment of minority group children: With bias toward none* (pp. 15–26). Paper presented at a national planning conference on nondiscriminatory assessment for handicapped children, Lexington, KY.
- Oakland, T., & Matuszek, P. (1976). Using tests in nondiscriminatory assessment. In T. Oakland (Ed.), *Non-biased assessment of minority group children: With bias toward none* (pp. 27–34). Paper presented at a national planning conference on nondiscriminatory assessment for handicapped children, Lexington, KY.
- Ortiz, S. O. (2001). Assessment of cognitive abilities in Hispanic children. *Seminars in Speech and Language*, *22*(1), 17–37.
- Ortiz, S. O., & Flanagan, D. P. (1998). *Gf-Gc cross-battery interpretation and selective cross-battery assessment: Considering referral concerns and the needs of culturally and linguistically diverse populations*. In K. S. McGrew & D. P. Flanagan (Eds.), *The intelligence test desk reference (ITDR): Gf-Gc cross-battery assessment*. Boston: Allyn & Bacon.
- Reschly, D. J., & Grimes, J. P. (1995). Best practices in intellectual assessment. In A. Thomas & J. P. Grimes (Eds.), *Best practices in school psychology III* (pp. 763–773). Washington, DC: National Association of School Psychologists.
- Salvia, J., & Ysseldyke, J. E. (1991). *Assessment* (5<sup>th</sup> ed.). New York: Houghton Mifflin.
- Samuda, R. J., Kong, S. L., Cummins, J., Pascual-Leone, J., & Lewis, J. (1991). *Assessment and placement of minority students*. New York: C. J. Hogrefe/Intercultural Social Sciences.
- Sandoval, J. (1979). The WISC-R and internal evidence of test bias with minority groups. *Journal of Consulting and Clinical Psychology*, *47*, 919–927.

Sandoval, J. (1998). Critical thinking in test interpretation. In J. Sandoval, C. L. Frisby, K. F. Geisinger, J. D. Scheuneman, & J. R. Grenier (Eds.), *Test interpretation and diversity: Achieving equity in assessment* (pp. 31–50). Washington, DC: American Psychological Association.

Sandoval, J. & Durán, R. P. (1998). Language. In J. Sandoval, C. L. Frisby, K. F. Geisinger, J. D. Scheuneman, & J. R. Grenier (Eds.), *Test interpretation and diversity: Achieving equity in assessment* (pp. 181–211). Washington, DC: American Psychological Association.

Sandoval, J., Frisby, C. L., Geisinger, K. F., Scheuneman, J. D., & Grenier, J. R. (Eds.). (1998). *Test interpretation and diversity: Achieving equity in assessment*. Washington, DC: American Psychological Association.

Sandoval, J., Zimmerman, I. L., & Woo-Sam, J. M. (1983). Cultural differences on WISC-R verbal items. *Journal of School Psychology, 21*, 49–55.

Sattler, J. (1992). *Assessment of children* (3rd ed.). San Diego, CA: Sattler.

Valdés, G., & Figueroa, R. A. (1994). *Bilingualism and testing: A special case of bias*. Norwood, NJ: Ablex.

Woodcock, R., & Muñoz-Sandoval, A. (1996). *The Bateria-R: Pruebas de habilidades cognitivas*. Chicago, IL: Riverside.

## ANNOTATED BIBLIOGRAPHY

Cummins, J. C. (1984). *Bilingual and special education: Issues in assessment and pedagogy*. Austin, TX: PRO-ED. A classic text in the area of bilingual assessment especially with respect to the issue of language proficiency.

Cummins discusses the concepts of language proficiency including Basic Interpersonal Communicative Skills and Cognitive Academic Language Proficiency that are critical to understanding the effect of language development on learning.

Flanagan, D. P. & Ortiz, S. O. (2001). *Essentials of cross-battery assessment*. New York: John Wiley.

In addition to providing a modern, theoretical framework for cognitive assessment, this book contains classifications of tests according to cultural loading and linguistic demand as well as a culture-language matrix that can assist practitioners in determining the relative influence of level of acculturation and English language proficiency on test performance. In addition, specific guidelines for nondiscriminatory interpretation of test results are provided, and the text is written specifically for practitioners in the field.

Sandoval, J., Frisby, C. L., Geisinger, K. F., Scheuneman, J. D., & Grenier, J. R. (Eds.). (1998). *Test interpretation and diversity: Achieving equity in assessment*. Washington, DC: American Psychological Association.

This is perhaps the most up-to-date and comprehensive volume on equitable assessment. The chapters are written by leading researchers who present a broad range of issues involving diversity with a balanced perspective and attention toward practical issues.

Valdés, G. & Figueroa, R. A. (1994). *Bilingualism and testing: A special case of bias*. Norwood, NJ: Ablex.

This volume is the definitive treatise on the issue of bias in the assessment of culturally and linguistically diverse children. Although not intended as a guide to assessment, the extensive review of empirical findings, judicial influences, and recommendations for assessment are crucial to the complete understanding of issues of bias.