Best Practices in Nondiscriminatory Assessment

Samuel O. Ortiz
St. John’s University

OVERVIEW

Nondiscriminatory assessment is often thought of as a set of methods and procedures that are applied in evaluations conducted on individuals from diverse cultural and racial backgrounds. In a similar vein, the phrase bilingual assessment has become the most common way of referencing evaluations on individuals from language backgrounds that are other than monolingual English. Of course, individuals who are from language backgrounds other than English frequently are culturally different as well as racially diverse. Thus, although each type of evaluation rightly encompasses unique issues, neither the former concept nor the latter phrase adequately captures the essence of nondiscriminatory assessment broadly speaking.

It is the intention of this chapter to provide a broad, comprehensive framework for assessment that may guide evaluation efforts independent of the particular issues of cultural, racial, linguistic, or other kinds of diversity that may present themselves. That is, the focus will be on providing an outline for systematic evaluation that promotes equity and justice across and throughout the entire process and not on techniques specific to one population or another. Other chapters in this edition of Best Practices are available to guide such assessment, and although various aspects of diversity are included here, the discussion centers on ensuring the creation of a reliable and valid process that promotes equitable outcomes and reduces, as much as possible, any potential discriminatory aspects of evaluation.

It would not be incorrect to state that nondiscriminatory assessment actually applies to all individuals, not just those who may be considered diverse or different. This is because diversity is much broader than just race, ethnicity, culture, or language. Indeed, nondiscriminatory assessment may be characterized simply as good assessment. However, because the procedures and tools used in assessment are all culturally bound, the potential discriminatory effects from their use with culturally (and other) different individuals is magnified and deserves special attention. This notion was recently reinforced in School Psychology: A Blueprint for Training and Practice III (Ysseldyke et al., 2006) where one of the more significant changes in the document, which outlines a vision for the future of school psychology, is the establishment of competency in all aspects of diversity as a foundational domain, one of four upon which all other domains of competence rest. The ability to manage issues of diversity effectively and competently within the context of assessment is predicated upon the application of a systematic, comprehensive nondiscriminatory framework.

Projections regarding the changing nature of the U.S. school-age population are consistent in citing a dramatic and relentless increase in the racial, ethnic, cultural, and linguistic diversity (National Center for Education Statistics, 2005; U.S. Census Bureau, 2000). In virtually every district in this country, school psychologists are faced with the prospect of evaluating the abilities, skills, and knowledge of students who come from unfamiliar cultural backgrounds and who speak languages not understood by the evaluators. Unfortunately, the manner in which cultural and linguistic factors combine to affect test performance continues to be under researched and poorly defined. As a result, the development of applied methods in assessment has not progressed adequately and has left school psychologists at a loss regarding the best way to approach evaluations of individuals from diverse backgrounds.
The reason for the seeming lack of progress in understanding the potentially discriminatory aspects of testing and the development of methods to account for it probably lies in the way in which test performance of diverse individuals has been viewed historically. When immigrants were first tested by Goddard (1913) at Ellis Island, intrinsic and hereditarian explanations for low scores were favored over consideration of the fact that they were not familiar with elements of the test that required knowledge of U.S. culture and that they knew little if any English. Even where data were relatively clear, such as those presented by Brigham (1923) from administration of the Binet scales to native English and nonnative English speakers, which demonstrated an increase in performance as a function of length of residence in the United States, there was a tendency to discount acculturation and limited English proficiency as confounding variables and instead point to genetic factors interacting with intelligence to align immigration patterns (Brigham, 1923). What emerged was the notion that bilingualism itself created some sort of handicap that was most easily remediated by learning English.

The concern with fair and equitable psychological assessment was not taken seriously until the very values of U.S. culture changed (Oakland & Laosa, 1976; Sandoval, 1998). As people began to protest discrimination in general and seek equal treatment under the law in the 1950s and 1960s, issues regarding the role that discriminatory assessment might play in creating various inequalities finally became a valid topic of interest (Irons & Guitton, 1993). Ultimately, the process of evaluation came under perhaps its greatest attack relative to the identification of children with disabilities (Jones, 1988; Laosa, 1976; Oakland & Laosa, 1976; Valdés & Figueroa, 1994). By the late 1960s, it had become evident that children from ethnic minority backgrounds were disproportionately represented in special education programs (Dunn, 1968), and following a slew of legal cases relevant to this issue in the early 1970s, passage of the landmark Public Law 94-142, the Education of Handicapped Children Act (later reauthorized and renamed the Individuals with Disabilities Education Act [IDEA]), set the first legal standards requiring evaluation practices to be fair, equitable, and nondiscriminatory. In its current incarnation, after the 2004 reauthorization, wording remains in IDEA that provides a clear directive regarding the need to engage in equitable 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” (Section 300.532).

Attention to the attenuating effects of racial, cultural, and linguistic differences on test performance is evident in this and other sections of IDEA 2004, including new and specific emphasis on distinguishing between intrinsic learning problems and problems that result as a function of limited English proficiency. School psychologists should take note that IDEA 2004 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).

To the present day, unbiased assessment remains an illusion. Complete elimination of bias is unrealistic and an impossible task. Thus, the goal of nondiscriminatory assessment should be viewed as an effort to reduce it to the maximum extent possible. Although many discrete methods for reducing bias in assessment have been presented (Flanagan & Ortiz, 2001; Gonzalez, Brusca-Vega, & Yawkey, 1997; Gopaul-McNicol & Thomas-Presswood, 1998; Ortiz & Ochoa, 2005; Plank, 2001; Rhode, Ochoa, & Ortiz, 2005; 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 is the 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 are likely to 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 (Notari-Syverson, Losardo, & Lim, 2003; Ochoa, 2003; Reynolds & Carson, 2005; Saenz & Huer, 2003; Sandoval et al., 1998; Vazquez-Nuttall et al., 2007). 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 (Ortiz, 2002; Ysseldyke et al., 2006).

BASIC CONSIDERATIONS

In recognition of the increasing diversity of the U.S. public school population, professional organizations have responded by setting new guidelines and standards for both the 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 (APA, 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). In 1999, additional guidelines were provided with the publication of The Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999), which outlines a clear professional responsibility for engaging in testing practices and using tests in a manner that reduces the potential discriminatory aspects of the process. As noted previously, Blueprint III is a document that provides a vision for the future of school psychology practice and which brought cultural competency in all areas of practice, including assessment, to the forefront and placed it as a foundational, not supplemental, aspect of service delivery.

The majority of attempts to develop procedures for nondiscriminatory assessment tend to revolve more around the use of tests than around the entire process. For example, cross-battery assessment and the Culture-Language Interpretive Matrix (Flanagan & Ortiz, 2001; Flanagan, Ortiz, & Alfonso, 2007; Ortiz, 2004; Ortiz & Ochoa, 2005; Plank, 2001), the Multidimensional Assessment Model for Bilingual Individuals (Ortiz & Ochoa, 2005; Rhodes et al., 2005), test modifications, use of interpreters, and dynamic assessment (Reynolds & Carson, 2005; Saenz & Huer, 2003; Vazquez-Nuttall et al., 2007), are all primarily concerned with bias that may result from the use of tests. 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 (Ochoa, 2003; Reynolds & Carson, 2005; Valencia & Suzuki, 2001).

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. Indeed, publication of the previous version of this chapter (Ortiz, 2002) represented the first comprehensive framework for nondiscriminatory assessment made generally available to school psychologists. This chapter continues to emphasize the need to engage in a wide variety of procedures that are designed to reduce bias in as many areas of practice as possible in a systematic manner. 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. In addition to the discussion of the main issues involved in nondiscriminatory assessment, the importance of applying a comprehensive and guiding model is reinforced by the delineation of best practices in the latter portion of this chapter, which continue to be the basis of a practical nondiscriminatory assessment framework.

Hypothesis Testing

Critical to nondiscriminatory assessment is the process 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 preconceived 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; Ochoa, 2003; 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 language comprehension, that may be related to the learning problem. Conversely, attributions of behavior that are made on the basis of stereotyped or preconceived 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). 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 that support the preconception 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; Ochoa, 2003; 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. Consider, for example, the use of techniques such as records review, interviews, observation, and testing–informal (i.e., RIOT), which are important components of any assessment and often taken as intrinsically valid. Yet, the very questions asked during interviews, the behaviors observed in the classroom, the work samples chosen for analysis, and expectations of performance on various tasks, will all be influenced by such preconceptions. Asking parents when their child first learned to walk 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 preexisting 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 to 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 underlying cause of poor performance. Even when there may be a diagnostic 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
training programs will need to embark on a focused
toward development of cultural competence from their
enough to receive direct experience and education
raised natively in a particular culture to derive such
Matsumi, 2006; Ortiz, 2006). Psychologists need not be
all assessment data (Lopez, 2006; Okazaki & Tanaka-
guide and context for collecting and evaluating any and
customs of a particular culture that can be used as both
experience with, the values, attitudes, beliefs, and
competence reflects a knowledge base of, or direct
Cultural and linguistic competency 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; Valencia & Suzuki, 2001). 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; Figueroa & Newsome, 2006; Frisby, 1998; Ochoa, 2003; Ortiz & Dynda, 2005; Rhodes et al., 2005; 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 (Ochoa, 2003; Ortiz & Ochoa, 2005; Reynolds & Carson, 2005; 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 individual 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

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 (Lopez, 2006; Okazaki & Tanaka-Matsumi, 2006; Ortiz, 2006). Psychologists need not be raised natively in a particular culture to derive such competence, but the necessary skills will not develop by reading a book or taking a trip. 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; Ortiz, 2006). In general, this process includes development of competence in three broad domains: knowledge, communication, and skills (Ortiz, 2006; Ysseldyke et al., 2006). 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.
second language (ESL) 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 (Lopez, 2006; Ortiz, 2006). 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 multilanguage 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 school psychologist 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 (Lopez, 2006; Ochoa, 2003; Reynolds & Carson, 2005; Vazquez-Nuttall et al., 2007). 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 (Figueroa & Newsome, 2006; Ortiz & Dynda, 2005; Ortiz & Ochoa, 2005). 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 (Ortiz & Ochoa, 2005; 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 (Reynolds & Carson, 2005). 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 (Figueroa & Newsome, 2006; Reynolds & Carson, 2005).

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; Reynolds & Carson, 2005; 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 (Reynolds & Carson, 2005; 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 was found (Jensen, 1974, 1980; Reynolds & Carson, 2005; Sandoval, 1979; 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, tests measure intelligence, or different facets of cognitive ability, or they predict achievement as well as if not better than anything else irrespective of race or ethnic origin (Neisser 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 aspira-
tions, 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 range of cultural content possessed by the individual, and incorrect inferences may be drawn (Ortiz & Dynda, 2005; Ortiz & Ochoa, 2005; Valdés & Figueroa, 1994; Vazquez-Nuttall et al., 2007). Tests may not be psychometrically biased, but they are culturally loaded and linguistically demanding to varying degrees (Ortiz & Ochoa, 2005; 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 from 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 in which 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 nonnative 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 has provided little insight into the performance of bilinguals. The emergence of native language tests with much better technical qualities (e.g., Bateria-III, Woodcock & Munoz-Sandoval, 2004; WISC-IV Spanish, Wechsler, 2005) may serve as the vehicles by which this distinct research question is addressed, depending on norm sample construction. For example, despite claims of fairness, tests like the WISC-IV Spanish continue to create norm samples that are still very different from the individuals on whom the test is likely to be used. Basing norms simply on whether an individual is or is not bilingual or on country of origin begins to address some of the issues but by no means represents the necessary level of attention to norm sample construction that would be required to create a test that would indeed be representative of the types of individuals one may encounter in any given assessment. Not only are bilinguals an extremely heterogeneous group themselves, cultural variations also go far beyond country of origin and must include attention to the nature and type of instruction received, family socioeconomic status, and parents’ level of education, to name a few (Ortiz & Dyna, 2005).

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 addition to the considerations that have been discussed, drawing valid and defensible inferences from assessment data requires that nondiscriminatory assessment practices be multifaceted 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 bias reduction in the various aspects of the assessment process. Based on the framework outlined in the previous version of this chapter (Ortiz, 2002), its various steps have been revised and refined and are summarized in Table 1. The framework combines 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

<table>
<thead>
<tr>
<th>Table 1. A Comprehensive Framework for Nondiscriminatory Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Assess for the purpose of intervention</td>
</tr>
<tr>
<td>• Assess initially with authentic and alternative procedures</td>
</tr>
<tr>
<td>• Assess and evaluate the learning ecology</td>
</tr>
<tr>
<td>• Assess and evaluate language proficiency</td>
</tr>
<tr>
<td>• Assess and evaluate opportunity for learning</td>
</tr>
<tr>
<td>• Assess and evaluate educationally relevant cultural and linguistic factors</td>
</tr>
<tr>
<td>• Evaluate, revise, and retest hypotheses</td>
</tr>
<tr>
<td>• Determine the need for and language(s) of formal assessment</td>
</tr>
<tr>
<td>• Reduce bias in traditional testing practices</td>
</tr>
<tr>
<td>• Support conclusions via data convergence and multiple indicators</td>
</tr>
</tbody>
</table>
Assess for the Purpose of Intervention

Assessment and evaluation activities are not interventions in and of themselves. Consider that when an individual has been assessed and it has been determined that no disability exists, the original problems that prompted the referral for evaluation 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. Some assessments are undertaken for the purpose of determining appropriate instruction or placement, irrespective of disability (e.g., within a three-tier model for differentiated instruction; see Tilly, chapter 2, vol. 1) or for exiting students from services (e.g., dismissal from special education; see Powell-Smith & Ball, chapter 15, vol. 2). It is abundantly clear that assessment of any kind is of little value unless it can be extended to incorporate appropriate interventions and treatment options irrespective of the identification of a disorder or disability. In school-based evaluations, modifications to the instructional program and the provision of specific remedial strategies are necessary whether or not the individual qualifies for special education services. Because a process of nondiscriminatory assessment will generate 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. It is for this reason that intervention-driven evaluation must necessarily be the first step in nondiscriminatory assessment. Failure to engage in a process that will generate data for the purpose of intervention can be construed as the most discriminatory aspect of assessment because it will bias all subsequent activities.

Assess Initially With Authentic and Alternative Assessment Procedures

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. The rapid and recent emergence of response to intervention (RTI) in school psychology attests to the value of intervention-based assessment and its potential value in reducing some of the discriminatory aspects of evaluation, particularly those procedures that do not focus on the appropriateness or adequacy of classroom instruction (Kovaleski & Prasse, 2004). Implementation of a proper three- or four-tier RTI framework is an example of a rigorous approach to utilizing authentic methods at this step (Tilly, chapter 2, vol. 1). Likewise, in addition to the activities that comprise a RIOT procedures (record review, interviews, observation, testing-informal), other examples of authentic or informal procedures include informal analysis of work samples, curriculum- or criterion-based assessment, performance-based assessment, portfolio assessment, and various test-teach-test frameworks such as dynamic assessment (Fischer & King, 1995; Lidz, 1997; Saenz & Huer, 2003; see also Batsche, Castillo, Dixon, & Forde, chapter 10, vol. 2; Upah, chapter 12, vol. 2; and Howell, Hosp, & Kurns, chapter 20, vol. 2).

When properly applied, authentic and alternative assessment strategies can provide valuable information, especially in school-based and subsequent special education evaluations. In educational settings, authentic assessment often utilizes material that is being provided through direct classroom instruction. Evaluation of learning and performance through use of the curriculum-based materials and content reflects an authentic nondiscriminatory approach because it seeks to measure that which the student has actually been taught. Accordingly, comprehensive nondiscriminatory assessment should include information and data obtained through such methods. As noted, examples include curriculum-based assessment or authentic measures of academic achievement and skill development but could also include performance-based assessment that evaluates more by task completion within context than by 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).

Evaluation of data from these methods may be dependent on additional procedures including analysis of differences in learning, examination of rates of progress or improvement in learning or growth rates, and evaluation of educational needs or requirements. The major difficulty here is in establishing the appropriate basis for comparison. It is not always easy to match students on variables that at the very least might include equivalent levels of language proficiency in English and the native language, equivalent amounts of schooling both in the United States and in the native country if any, the type and length of educational programming received previously and currently, the student’s and family’s level of acculturation, and socioeconomic status, to name a few.

**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. For example, others might include health, family situations, socioeconomic issues, teacher biases, and ineffective instruction. Therefore, 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, nonacademic 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; interviews 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 differences should be viewed as examination of relatively unique circumstances or learning experiences that cannot 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 those 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.

**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; Ochoa & Ortiz, 2005; Rhodes et al., 2005). 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. The data may already exist in school districts where there are bilingual or 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; Saenz & Huer, 2003). 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. Specific strategies for working effectively with English learners in the course of assessment are beyond the scope of this chapter, and the reader may wish to consult other sources for detailed information (e.g., Ochoa & Ortiz, 2005; Ortiz & Dynda, 2005; Ortiz & Ochoa, 2005, Rhodes et al., 2005).

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 and 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; parents’ 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, and expectations; system attitude regarding dual language learners; and socialization with peers versus isolation from peers.

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 those 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 development 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, the individual’s birth order or relative impact of siblings and his or her language development, the individual’s fluency in the native language and English, individual’s and parents’ level of acculturation, parents’ fluency in the native language and English, parents’ level of literacy in the native language and English, parents’ level of education, and parents’ socioeconomic status.

**Evaluate, Revise, and Retest 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, supported by a convergence of data and multiple information, 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, but also to bolster conclusions and inferences that are put forth from analysis of all available data. Additionally, there should also be evidence that systematic and appropriate efforts to improve the student’s classroom performance, for example, a three- or four-tiered RTI model, were undertaken but proved to be unsuccessful. The limus 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 problems. When difficulties learning in the classroom setting or behavioral problems cannot 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.

**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 2004 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 (Sections 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 prerelief data, the unique background variables of the individual, and relevant referral concerns. Because each case is unique, and because the foundation of IDEA 2004 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 prereferral intervention and assistance process. It seems likely that the advent of RTI procedures will provide increased attention to the process that may assist practitioners in gathering data during this phase. Although RTI does not necessarily result in unbiased data collection, it certainly allows for the collection of data and information that may inform decisions related to distinguishing difference versus disorder long before the use of tests is contemplated. Use of a prereferral 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 prereferral 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.

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 et al., 2007; Ortiz & Ochoa, 2005; Ortiz & Dynda, 2005; Plank, 2001; Rhodes et al., 2005; Vazquez-Nuttall et al., 2007). 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 Culture–Language Interpretive Matrix (Flanagan & Ortiz, 2001; Flanagan et al., 2007; Ortiz, 2001; Rhodes et al., 2005) 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 (Notari-Syverson et al., 2003; Reynolds & Carson, 2005; Saenz & Huer, 2003; Valencia & Suzuki, 2001). 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 nonstandardized ways may rest more in the collection of qualitative versus quantitative data.

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 nonverbal 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 information is often much more useful in instructional planning than any quantitative results that might be 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.

Because of the violations of standardization, it is perhaps most prudent that such protocols not be scored and no quantitative data reported so that the tendency to interpret and assign meaning to the numbers is obviated. Moreover, use of tests in a qualitative way should remain guided by efforts to intervene and not simply to diagnosis. School psychologists who have the requisite bilingual skills to conduct assessment in the native language must recognize the limitations of the process and look more for valuable information about functioning that can only be gathered by such bilingual interactions and not by test scores.

**Support Conclusions Via Data Convergence and Multiple Indicators**

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.

A necessary 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. In some cases, it may be difficult to find convergence, for example, in cases where data are based on third-party ratings of behavior. Whereas data from standardized tests are expected to converge, for various reasons data from independent ratings of psychopathology may not, leaving the school psychologist to decide which data should be ascribed the most credibility. In practice, a preponderance of evidence is sufficient to provide validity to conclusions, but 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. In the final analysis, nondiscriminatory assessment holds that equivocal data supports the null hypothesis, that functioning is within normal limits, and that any observed difficulties are the result of factors other than internal disability.

**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 a 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 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 a test but rather differences in experience between an individual taking the test and the individuals on whom the test was normed. However, 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 that assists not only in interpreting data fairly but also in 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 that 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 prereferral activities. These activities provide data that informs 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 essential to nondiscriminatory assessment can be 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 is increased.

REFERENCES


ANNOTATED BIBLIOGRAPHY


Updates the modern, theoretical framework for cognitive assessment originally found in the first edition and expands development of the classification of tests according to cultural loading and linguistic demand and provides a clearer operationalization of the Culture–Language Interpretive Matrix. The matrix assists practitioners in determining the relative influence of level of acculturation and English language proficiency on test performance, simplifying the task of delineating difference from disorder. Specific guidelines for nondiscriminatory interpretation of test results are provided. Written specifically for practitioners in the field.


Discusses specific aspects of assessment with diverse individuals including language, academic, and cognitive evaluation. Covers the full range of comprehensive nondiscriminatory assessment and can serve as both a resource for practitioners as well as a solid text for teaching. Should be considered the first source for psychologists who wish to learn more and develop better competency in assessment of diverse individuals.

Remains an important, comprehensive, and well-written volume on equitable assessment. Chapters are written by leading researchers who present a broad range of issues involving diversity with a balanced perspective and attention toward practical issues.


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.