The Psychoeducational Assessment of Ebonics Speakers: Issues and Challenges

Sharon-ann Gopaul-McNicol, Grace Reid, and Cecilia Wisdom, Howard University

Traditional standardized psychoeducational assessments do not adequately consider, nor do they fully account for or adapt to, the nonstandard dialects and cultural experiences that certain test takers bring to the evaluation experience. The assessment of African American Ebonics-speaking students using such tests presents particular challenges. This article focuses on the specific limitations of these assessments for Ebonics speakers, and describes alternative measures that yield more accurate results for these students. Further, the article highlights the implications of traditional and nontraditional assessment approaches for psychoeducational test developers, evaluators, educators, and students.

INTRODUCTION

Issues surrounding the psychoeducational assessment of bilingual/bidialectal children have drawn considerable attention from educators, psychologists, researchers, and politicians (Armour-Thomas & Gopaul-McNicol, 1997; Mercer, 1979; Oakland, 1977; Vasquez-Nuttall, Goldman, & Landurand, 1983; Williams, 1975). It has been widely asserted in recent decades that traditional assessments of cognitive ability and written, oral language, and reading skills do not yield accurate results for African American children who speak nonstandard dialects of English such as Ebonics. Several scholars have claimed that the test-performance disparities noted between students who speak Standard English and those who speak Ebonics can be attributed to dialectal as well as cultural differences (Allen & Boykin, 1992; Boykin, 1994; Gordon, 1991; Hilliard, 1983; Jones, 1988). Hilliard further concludes that Black children’s poor performance on these assessments often results in their being labeled “poor readers, low in intelligence, and speech impaired” (p. 52). Moreover, errors wrought by the “indiscriminate use of psychological tests, especially IQ tests” have been found to contribute to the disproportionate number of Ebonics-speaking African American children who are relegated to special education classes (Cummens, 1984, p. 1).

Hilliard and others (Baugh, 1991; Haynes & Gebreyesus, 1992) have asserted that erroneous assumptions about the inferiority of the language systems and culture of African Americans vis-à-vis that of European Americans are at the core of the tests’ inaccuracies. Chief among these misconceptions, according to Hilliard, are those which attest that Standard English is a pure and superior language, that English usage is fixed and permanent, and that American English is uninfluenced by African languages. Additionally, many educators and psychoeducational evaluators are unaware of certain basic facts about the distinctive African American language system called Ebonics (see other articles in this issue: Harper, 1999; Taylor, 1999; Wright, 1999; see also Williams, 1975). As a result,
they subsequently fail to understand Ebonics in terms of its multiple linguistic origins, particularly those features resulting from the fusion of various African languages with English. They also fail to consider the enduring and far-reaching impact of Ebonics on African American communication, comprehension, and interpretation generally or, more specifically, on the education of African Americans.

How are evaluators, teachers, and test developers to contend with these issues? In seeking to answer this question, the present article examines the impact of dialectal differences on the performance of Ebonics-speaking and Standard English-speaking test takers on traditional standardized psychoeducational assessments. It also describes alternative assessments that have been shown to facilitate more accurate portraits of the psychoeducational functioning and potential of Ebonics speakers.

**Ebonics and Traditional Assessments of Psychoeducational Functioning**

Garner and Rubin (1986) and Roy (1987) emphasize that every language is shaped by the cultural experiences of the social group that uses it, and that no one language is better than any other insofar as its structure is concerned. However, dialectal differences, particularly those in phonology and grammar, often correlate with ethnicity, geographic location, and socioeconomic status (Berko Gleason, 1997; Homel, Palif, & Aaronson, 1987; Roy, 1987; Speicher & McMahon, 1992; Warren & McCloskey, 1997). Baratz and Shuy (1969) and Shuy (1982) offer several examples of grammatical differences in the language patterns of White speakers of Standard English and Black speakers of Ebonics, noting that evaluators often view Ebonics usage as representative of vocabulary deficiencies or even as symptomatic of impaired auditory functioning.

Because language is typically used as a vehicle for securing information on any test, irrespective of the content area, dialectal differences that affect language comprehension exert strong influences on test results. For most of the traditional psychoeducational tests (i.e., the Weschler Intelligence Scale for Children–III, the Weschler Intelligence Achievement Test, the Wide-Range Achievement Test) used in classrooms in the United States today, proficiency in Standard English is a basic criterion. That is, these tests are constructed such that the normative or correct response is most often the one that reflects proficiency in Standard English. Answers that deviate from Standard English are considered incorrect, even if they conform to the rules of a language system utilized by millions of people. Clearly then, the more limited a language assessment test is in its ability to take into consideration diverse varieties of language, the more likely that it will be inappropriate for test takers who are not part of the population upon which it was normed (Armour–Thomas & Gopaul–McNicol, 1998). Under such conditions, it is also more likely for evaluators to conclude that students who speak nonstandard forms of English have poor or substandard verbal abilities or are language-delayed (Baratz & Shuy, 1969).

The purport of standardized psychoeducational achievement tests is to measure the knowledge that students have acquired from their formal education. Paradoxically, for speakers of Standard English in the U.S., such tests actually measure grammatical knowledge that in all probability has been acquired from and reinforced by familial and community influences. It is only for students from communities in which the use of vernacular dialects such as Ebonics is prevalent that the tests measure knowledge that has been acquired primarily in school and that is not subsequently reinforced at home or in the community. However, as Wolfram (1990) has noted, the distance between students' everyday language system and the language system used on standardized psychoeducational assessments is directly proportional to the likelihood that task interference due to language differences will negatively affect those students' test performance.
For example, children who use Ebonics often encounter words on traditional psycho-
educational assessments that they are prone to confuse with other words that, because
of dialectal rules, are not differentiated in their pronunciation. Two pairs of words that
can be confused in this manner are “mild” and “mile” (due to a phonological rule in
Ebonics that calls for the elimination of terminal consonants) and “death” and “deaf”
(which are both pronounced “deaf,” due to a dialectal replacement of the phoneme “th”
with “f” in Ebonics). Other pairs subject to confusion are “axe” for “ask” and “scrimp”
for “shrimp,” with the first words in each representing the Ebonics pronunciation of the
latter. When children whose spoken dialect is Ebonics see or hear such words as part of
an assessment exercise, they are likely to confuse the Standard English form for the
Ebonics one, or vice-versa (Lee, 1994; Pietsrup, 1973; Wolfram, 1979). Examiners who are
unfamiliar with the dialectal differences between Ebonics and Standard English must
learn how to determine in which language system Ebonics-speaking children’s responses
are based in order to accurately ascertain their language comprehension skills—and
ultimately develop fairer, more accurate assessments.

Assessment challenges arising from dialectal differences are not limited to pronuncia-
tion, grammar, and vocabulary—even though these are the aspects most commonly noted
in discussions of this topic. Many traditional assessments of language comprehension
such as the Weschler scales are based on the extension of language beyond literal meanings.
Such measures assess children’s abilities to draw inferences, which invariably are informed
by their background knowledge, which is first and foremost determined by their cultural
frames of reference. This accounts for the differing degrees of familiarity that speakers
of Ebonics and speakers of Standard English have with vocabulary items that typically
appear on traditional psychoeducational tests. Vocabulary items that are central to the
culture of the latter may be peripheral to or even absent from that of the former. Moreover,
the subjects to which Ebonics-speaking children are exposed in their home and community
environments, as well as the cultural values associated with those subjects, create the
frame of reference that they bring to testing situations. Thus, cultural differences can be
viewed as affecting children’s scores on inference-based tests (Armour-Thomas, 1992).

Given these dialectal and cultural differences, it is simply inappropriate to compare
the test scores of Ebonics speakers with those of Standard English speakers as if both
groups start from the same linguistic baseline. The language system of Ebonics differs so
dramatically from the Standard English emphasized on most of the psychoeducational
assessments used in U.S. schools today that Ebonics-speaking students often do poorly
on these tests. Moreover, Ebonics speakers are not part of the population on which these
assessment instruments were normed; thus, the unsuitability of such instruments for these
students is magnified.

**Alternative Psychoeducational Assessments for Speakers of Ebonics**

**Contextualized Assessments**

Traditional tests of language comprehension (i.e., the Weschler Intelligence Scale for
Children–III, the Woodcock–Johnson Test of Cognitive Ability) require children to define
words that are presented in isolation. Typically, these tests do not take into account that
language skills are largely acquired by listening to everyday speech rather than through
formal instruction, and that many words have different meanings in different cultures.
The word, “tostone,” for instance, means “quarter” or “half dollar” to a Chicano, whereas
to a Puerto Rican it refers to a portion of a fried banana. Similarly, as noted earlier, the
words, “axe” and “scrimp” have different meanings for Ebonics speakers than for speakers
of Standard English. For this reason, on tests that measure verbal proficiency, Ebonics speakers (as well as speakers of other nonstandard varieties of English) should be asked to provide (either written or verbal) sentence- or phrase-length responses to indicate whether they are providing Ebonics or Standard English responses to vocabulary items (Armour–Thomas & Gopaul–McNicol, 1997, 1998). Once it has been determined that Ebonics speakers consistently provide incorrect answers to specific items—or rather, that they consistently provide responses that are correct for nonstandard forms of various words but not correct for the normative forms—then those items should be reframed or eliminated from test batteries. This is not recommended to reinforce continued misunderstanding or to “dumb down” the tests, but simply to facilitate a fairer consideration of dialectal differences and a more accurate assessment of the language comprehension and knowledge of Ebonics speakers.

The finding that an individual performs a given task very well or very poorly in one context does not guarantee that he or she will or will not be able to perform a similar task in a different context. For example, several researchers have found that a competency in arithmetic for the purpose of carrying out everyday tasks does not necessarily correlate with success on standardized arithmetic tests, even if both the tasks and the tests require the use of the same arithmetic operations (Lave, 1977; Murtaugh, 1985; Rogoff, 1978). Carraher, Carraher, and Schliemann (1985) note that Brazilian “street children” who intuitively developed models of probability in order to serve as street brokers of lottery tickets, had difficulty applying these models to similar types of probability problems in formal educational environments. Given the effects of environmental factors on educational performance and potential, a more qualitative approach to assessment is warranted for students who speak Ebonics. Evaluators must come to understand the forms that various cognitive skills take when manifested in these students’ cultural environments, and subsequently determine which of these forms are analogous to skills that traditional assessments require students to demonstrate. Having made such determinations, examiners can then more accurately assess the skills and potentials of Ebonics speakers.

For example, Sternberg (1986) found that the performance gap between subjects from two cultural groups (Venezuelan and North American) narrowed or even disappeared when members of the groups were administered assessments that (a) used culturally (and linguistically) appropriate materials, (b) provided revised instructions replete with contextualized terms, (c) were administered by examiners who were sensitive to the nuances of both groups’ cultural and dialectal differences, and (d) were scored such that culturally analogous cognitive abilities were recognized.1 Psychoeducational evaluators who consider themselves more than just statistically minded psychometricians will find these approaches indispensable.

Ecological Assessments

Testing professionals can turn to a battery of other contextualized assessment tools to achieve accurate portraits of Ebonics-speaking students’ performance and ability, including interviews and questionnaires designed to ascertain their home language, language samples (ethnographic information on test takers’ language use), and observations in

---

1. A caveat, however, is that the statistical validity of a standardized test will most likely be reduced when modifications are made to either the test’s content or to its administration procedures. Data obtained through such assessment measures must therefore be interpreted qualitatively rather than quantitatively—that is, examinees’ performance of various tasks should be evaluated in terms of their strengths and weaknesses. Nevertheless, Sternberg asserts that an immense amount of otherwise unobtainable information can be acquired by administering culturally equivalent tests.
real-life contexts (Gopaul–McNicol & Armour–Thomas, 1997; Lopez, 1995; Payan, 1989; Prutting, 1983). The information gained from such assessments of students in their most familiar environments—that is, in their homes and communities—has been shown to broaden examiners’ understanding of diverse linguistic competencies (Gopaul–McNicol, 1992; Hamayan & Damico, 1991; Schiff–Myers, Dukic, McGovern–Lawler, & Perez, 1994). When using these approaches, however, these researchers note that examiners must be especially attentive to several factors and conditions, including the ways in which Ebonics-speaking children socialize and the activities in which they engage, the roles that these students play in social situations, and the respect (or lack of respect) they give to and are given by those in their home- and community-based settings (i.e., family, peers, friends).

The following case study illustrates the negative results that traditional psychoeducational assessments often yield for Ebonics-speaking students. It also reveals the manner in which such students, when observed in their most familiar environments, are often shown to evidence skills analogous to those on traditional assessments. This approach is recommended by the American Psychological Association (APA) (1993) in its guidelines for conducting comprehensive and qualitative ecological assessments of students from diverse cultures in settings outside of the school. For such students, the APA asserts that home and community visits—whereby contact is established with families, community leaders, church representatives, and so forth—are critical if evaluators are to accurately assess students’ true cognitive achievement and potential.

**BACKGROUND:** Aisha (a pseudonym) is a 14-year-old African American girl in the eighth grade who speaks Ebonics and Standard English alternately. Her teacher’s records indicated that Aisha was functioning at a fourth-grade level in mathematics and a third-grade level in reading. The teacher felt that a special education program would better address Aisha’s academic delays, so she referred Aisha to a psychoeducational evaluator to ascertain which program would be appropriate.

**TRADITIONAL ASSESSMENT RESULTS:** On the Wechsler Intelligence Scale for Children–III (WISC–III), Aisha obtained a full-scale IQ score that placed her in the moderate range of retardation in both the verbal and nonverbal areas. Individual subtests indicated moderate retardation and cognitive delays in all areas assessed. This assessment designated Aisha as eligible for a full-time special education program of modified institutional services (MIS-I).

**ECOLOGICAL ASSESSMENT RESULTS:** Alternative, non-language-based assessments provided an entirely different picture of Aisha’s cognitive abilities and academic potential than did the traditional Standard English-based WISC–III assessment. Observations and interviews conducted in Aisha’s home and in her informal community settings revealed that in those contexts she was considered smart and on par with the competencies expected of her peer group. Members of her family indicated that she often cooks, helps with the grocery shopping, assists in the care of her ailing great-grandfather, and fixes most of the family’s small appliances when they malfunction. Her language ability was found to be more borderline than deficient when more ecologically sensitive assessment techniques were utilized. Observations of Aisha at the neighborhood playground revealed that she was able to express herself verbally in a fairly capable manner, albeit in Ebonics. In her natural setting, she utilized words such as “dangerous,” which she had been unable to define on the written IQ test. Whereas Aisha had been unable to recall as many as seven numbers on the Digit Span subtest of the Wechsler scales, she could remember all the items on a 10-item grocery list with ease. Equally notable was her ability to perform calculations in basic addition and subtraction in the grocery store, even though she demonstrated no such mathematical aptitude on the WISC–III. Thus, overall, the results of ecological evaluation indicated that Aisha possessed good short-term memory, planning, perceptual, organization, and visual–motor coordination skills, and at least some competency in mathematics—abilities that are clearly incommensurate with moderate retardation. When she was tested using this approach, and when she was not penalized for speaking in and filtering her language comprehension through Ebonics, evidence of mental retardation for this child was unfounded.

**Curriculum-Based Assessments**

Another means by which the negative effects of dialectal differences on psychoeducational testing can be minimized is the use of assessments that are based on the instructional
materials and content actually covered in students' classrooms. Such evaluations often yield more accurate depictions of Ebonics-speaking children's strengths and weaknesses than can be derived from traditional standardized psychoeducational tests (Conoley & Conoley, 1992). Informal reading inventories (assessments of student reading from in-class and homework materials), error analyses, and task analyses are additional examples of curriculum-based assessments.

**Other Recommendations**

To reduce linguistic and cultural biases on standardized psychoeducational assessments, Vaughn–Cooke (1983) offers the following suggestions to language pathologists, test developers, and school psychologists:

(1) include speakers of vernacular dialects in standardization samples used to norm assessments;
(2) become familiar with the linguistic characteristics of the diverse communities from which students in such samples come;
(3) become more conscious of and sensitive to the cultural experiences of speakers of nonstandard dialects;
(4) learn to identify those responses made during assessment that are attributable to dialectal differences and therefore ought not to be regarded as "incorrect" and for which adjustments should be made; and
(5) develop new tests specifically designed to measure the intellectual potential of speakers of vernacular dialects.

Wolfram (1990) recommends that educators and other test administrators ask themselves several questions regarding the use of standardized assessments with students who speak nonstandard vernaculars of English such as Ebonics. Among these are the following:

(1) Does the test actually measure what it claims to measure?
(2) What assumptions about language did the test developers make when creating the test?
(3) What kinds of language-related tasks are demanded of test takers on the test?
(4) Do these tasks test the same ability for students of all backgrounds?
(5) Which linguistic and cultural groups (according to the demographic information provided in the test manuals) were used to “standardize” the test?
(6) How should the test results of speakers of nonstandard dialects be interpreted?

Lopez, Lamar, and Scully–DeMartini (1997) and Mattes and Omark (1991) offer some additional recommendations for improving the assessment of limited-English-proficient (LEP) students that are relevant to the assessment of Ebonics speakers. These include:

(1) the conduct of pre-referral activities such as diagnoses to determine if these students need to be assessed differentially in the first place;
(2) the determination of language dominance and proficiency prior to testing;\(^2\);
(3) the training of assessment personnel to incorporate more comprehensive and ecologically sensitive assessment approaches;
(4) the use of multiple assessment tools such as Armour–Thomas and Gopaul–McNicol’s (1998) Bio-cultural Assessment System;

---

\(^2\)The New York State Department of Education (1997) has recommended distributing the "Home Language Questionnaire," currently in use in all schools in New York City, to students in schools throughout the state. This instrument helps school officials assess students' proficiency in Standard English to determine whether any ancillary language development services are needed. Measures like this constitute the first of many steps that must be taken to address the effects that cultural and dialectal differences have on the psychoeducational assessment of bilingual/bidialectal students.
(5) the suspension or extension of time limits for testing and the provision of immediate feedback to LEP and non-Standard-English-proficient test takers.

Additionally, in the school classroom prior to assessment, Lopez and Gopaul-McNicol (1997) emphasize that teachers and examiners should avoid the continuous correction of speech and writing errors based on dialectal differences. As Roy (1987) maintains, such “spot corrections” do not help children learn Standard English, but rather increase students’ frustration and lower their self-confidence. Educators can also help reduce the frequency and severity of psychoeducational misdiagnoses by teaching and encouraging Ebonics speakers to use code switching, or the technique of variously employing Ebonics or Standard English, as appropriate, in different contexts. Code switching necessitates self-monitoring and self-evaluation, competencies that support the test-taking skills of Ebonics-speaking and other children. Thus, on the playground or in other informal settings, Cummins (1984) suggests that Ebonics-speaking students should be allowed to freely and confidently use their dialect. However, in formal learning settings, they should be expected to switch to the required, though not superior, Standard English dialect.

CONCLUSION

Traditional psychoeducational assessments do not adequately consider, nor do they fully account for or adapt to, the nonstandard dialects and cultural experiences that certain test takers bring to the evaluation experience. Given that psychoeducational functioning is labile and responsive to shifts in context, most of these assessments provide inaccurate portraits of these individuals’ language and other cognitive competencies. Although some members of the educational community function as if the literature on how dialectal and cultural differences affect psychoeducational assessment is extensive, continued research efforts are needed if the assessment of all students is to be truly fair and accurate.

The testing industry as a whole must reconsider some of its major assumptions about standardized testing, including the tendency to assume that all test takers share identical cultural frames of reference and linguistic backgrounds that can be measured by the administration of a single testing instrument. The development and validation of alternative models of assessment for Ebonics-speaking children is particularly critical, given these children’s traditionally low scores on traditional standardized assessments. Test developers need to create—and evaluators need to employ—a broader menu of assessment options to enable them to better recognize the many varieties of human talent that exist in our society. Further, psychoeducational evaluators are urged to question the use of the traditional psychoeducational linguistic battery in the assessment of children from diverse linguistic backgrounds, and to implement alternative assessment measures that better serve the needs of this population whenever possible. One question that must always be kept in mind is whether the tasks on standardized psychoeducational tests are relevant to the values held by the school communities. Psychoeducational evaluators-in-training must be formally introduced to the available research on bidialectism and the efficacy of programs that have resulted from this research. To expect them to arrive at empirically valid taxonomies of difference on their own, without exposure to the subject in an academic setting, is unrealistic. Likewise, in training future teachers, teacher educators must be clear about both the importance and role of Ebonics in the lives of those who speak it.

Finding an accurate means of assessing the psychoeducational achievement and potential of Ebonics speakers in educational and clinical settings remains one of the challenges facing the U.S. educational community today. Test developers, testing policymakers, evaluators, and educators alike must gain increased sensitivity to the needs of children who speak Ebonics—or any other dialect, for that matter—and they must work together to address this challenge.
REFERENCES


