

## SECTION II

# Language Issues in Assessment

All languages have structure, and all people use structure when speaking. No language can be superior or inferior, only different in the sense that it fits the needs of different groups. A child coming to school may speak a native tongue that is quite different from the standard English generally used in American schools. These differences may go well beyond simple dictionary definitions or diverse syntactical structures. They can affect the impressions youngsters give their teachers about their intellectual abilities, their attitudes, and even their social acceptability.

Wilma S. Longstreet, *Aspects of Ethnicity: Understanding Differences in Pluralistic Classrooms*, 1978.



As stated previously, assessment is a cultural phenomenon. The ways students are assessed in U.S. schools are by no means universal. Assessment is also in most instances a linguistic process. Assessment is dependent on language for giving instructions, for framing actual prompts or test items, and very often for the actual responses students are to give. Even so-called “nonverbal” tests, such as the Raven’s Progressive Matrices or other visual-spatial tasks, often require that students have the mental language to conceptualize the problem or hold certain ideas in memory (Oller, 1992; Roth, 1978).

Language is intimately tied to culture, and separating them is quite artificial. Language was actually discussed to a considerable extent in Section I, in terms of how students participate in conversation or discussion in the classroom, culture-based conventions for posing and asking questions, and cultural variations in using language to show what one has learned. In this section, there is an exploration of assessment of English language learners, and a discussion of the parallels between English language learners and dialect-learners. There are also a limited number of examples related to gender, showing how language use in assessment may affect males and females differently.

## How Culture Shapes Language Use: Implications for Assessment

Depending on their backgrounds, students may be inclined to communicate in ways that are quite similar to what’s expected in school or, conversely, quite different. More economically privileged families, for example, tend to use language in the ways schools demand, so their children will likely be more familiar with it. In particular, middle-class parents often ask their children to explain past events or talk about upcoming events. This “decontextualized” language use, in which the immediate topic of conversation is not present or self-explanatory, but exists in the past or future, has been associated with success in reading

(Snow, 1983). After all, to comprehend a text, the child needs to be able to rely on language alone. There is no immediate, real-life context to clarify the meaning of a sentence, as in face-to-face communication.

In another example, consider that some cultures equate wisdom with speaking very selectively and value listening over speaking (Philips, 1983), while other cultures, such as the dominant U.S. culture, equate power and knowledge with the active use of language in social situations (Greenfield, Quiroz, & Raeff, 1998). The latter, of course, more closely reflects what's expected in most classrooms. In similar contrast, some groups use language with children primarily to socialize them to expected behaviors, while other groups attempt to engage small children in talking about what they observe or experience — again, not unlike how language is used in schools (Snow, 1983). These very different ways of socializing and interacting with children can be expected to engender different skills — some matched to school expectations, some not.

It's not uncommon to make assumptions about people based on how they communicate. As educators, we need to be continually vigilant that our assumptions about our students are accurate. The following vignettes are just two examples of how students can be misperceived when teachers and other educators do not understand — or even recognize — that they have a different, culturally based, communications style.



#### VIGNETTE: SIGNS OF SUCCESS

*A small group of students from Alaska's St. Lawrence Island take part in a national competition that involves designing and constructing a model to solve an environmental problem. During the competition, teams from many other states are seen animatedly discussing the problem and debating strategies. The St. Lawrence team members, however, seem to talk very little among themselves. Instead, as they construct their model, team members contribute quietly to the work. To observers, little appears to be happening. Many of them conclude that the team has stalled, is bored, and isn't likely to generate anything of interest. But when the time comes for competition participants to show their work, the solution presented by the St. Lawrence students surprises the judges with its elegance and power.*



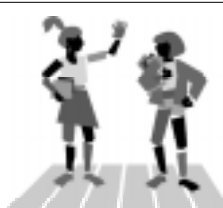
## THINGS TO CONSIDER

What strengths contributed to the success of these students? Until the end, competition observers did not see the social-interaction indicators they had traditionally equated with success. Teachers also expect certain indicators to provide evidence of students learning. We must take care not to misinterpret student work when their ways of learning do not match what we expect to see. The cognitive strengths of the Alaskan students were hidden by their reliance on nonverbal communication, and those observing the team were limited by their own cultural experiences and perceptions. The absence of overt verbal communication among team members was interpreted by the observers as limited interaction and enthusiasm. It is clear that the observers were wrong. But what *was* behind the team's apparent lack of communication?

---

Anthropologist and intercultural communication researcher Edward T. Hall (1977) has studied communication differences across cultures. His studies of the ways in which culture conditions us to perceive our world, and the others in it, provide important insights into ways we can, at times, “not see” or “not hear” what another is communicating.

One of the differences Hall has identified has to do with the amount of explicit information people provide when communicating about a topic. How much do they specifically explain about any given subject? In some cultures, members do not provide much elaboration; they don't



### DID YOU KNOW?

- High-context cultures tend to use communication strategies in which most of the meaning is embedded in physical and nonverbal means.
- Low-context cultures, by contrast, tend to use communications in which meaning is made explicit through details and elaborate verbal or written messages.

need to because they share extensive common experiences and values. This kind of communication is referred to as *high context*. By contrast, if the experiences and values within a group are far more diverse, communication *requires* more explicit elaboration and detail. This style is known as *low-context* communication. In all likelihood, the St. Lawrence Island community has a norm of high-context communication, as reflected in the team's behavior.



#### VIGNETTE: DARREN'S STORY

*Darren, an African American, is new to a second-grade classroom composed largely of European American students. His teacher (also European American) has divided her class into heterogeneous reading groups, and she randomly assigns him to one. During the meeting time for his reading group, she attempts to determine his reading level. She says, "Darren, would you like to read the first page?" He shakes his head. She is surprised because students don't usually refuse, but she lets it go. Other students read, and then she asks him a second time, "Darren, why don't you read the next page?" He shakes his head again. She looks at him sternly, wondering if he is going to be a discipline problem, or whether his reading skills are so poor that he can't manage the text. "Darren, please read the next page now." Darren complies and reads fluently, with a clear voice. The book actually appears to be way too easy for him. The teacher is confused by what has just occurred.*



#### THINGS TO CONSIDER

When Darren doesn't respond as she expects him to, his teacher initially worries that he might have a bad attitude or poor reading skills. In reality, any number of things may have been going on. Darren may have been shy or nervous about reading in front of a group on his first day at a new school. Because the teacher

phrased her requests as questions, he may also have thought that he had a choice about responding. Many African American students are accustomed to hearing commands in the form of assertive statements rather than as questions (or indirect requests) at home. Teachers who do not realize this often mistake a student's noncompliance as a negative attitude or lack of knowledge/competence.

---

## Different Oral and Literate Traditions

Just as different groups may use different dialects of the same language, so too they may differ in other aspects of how they use language and communicate meaning. There are differences that have been called “paralinguistic,” because they are not directly part of the language code. For example, students may use different kinds of intonation and levels of volume as they speak to one another. Are there highs and lows, or do voices sound somewhat monotone? Do they “sound” excited? Agitated? Uninterested? It is remarkable how we unconsciously assume interest level or emotional content from these paralinguistic features. For instance, Navajo speakers tend to use much lower volume and less variation in pitch, sounding monotone to the inexperienced listener. But these features cannot fairly be interpreted to mean lack of engagement with a topic or lack of emotion.

Speakers differ also in the use of gestures, how near they stand to one another, whether they focus directly on “the point” or “talk around it,” and numerous other features. All these differences in the ways we communicate what we know and who we are, as well as what we care about, are deeply rooted in our linguistic/cultural identity.

Patterns of text organization differ from culture to culture. A typical Western European and American pattern of argument is to state a thesis (e.g., via a topic sentence) and then support that thesis with details (evidence). When a student begins (orally or in writing) with a series of details and builds up to a conclusion, we may perceive that his/her strategy is immature or deviant. In fact, it is an equally logical strategy common in many other cultures.



### DID YOU KNOW?

Native people, unlike middle-class European Americans, are likely to tell a story to children without asking them to offer their interpretations of events as the story is told. Native children are expected to make their own sense of the tale without the kind of discussion or questioning common in classrooms. Through years of hearing a story repeated, their interpretations will steadily mature, without this discussion.

For many Pacific Island cultures, storytelling is a primary means of communicating cultural values, the history and relationships of family, important knowledge of natural phenomena, moral lessons, and more. In this tradition of family storytelling, stories do not end. Retellings add information, characters, and things to learn. Different storytellers may introduce different points of view about the same event. Each session is intended to provide a piece of the whole picture and leave listeners ready for another telling.

Teachers need to be aware of the different possible ways of relating information or telling stories, if they are to fully understand children from nondominant cultural backgrounds. It is important for teachers of diverse student populations to know something about “sociolinguistics,” the science of the social aspects of language. Otherwise, differences in children’s language can lead teachers to unintended biases about what children are capable of learning.

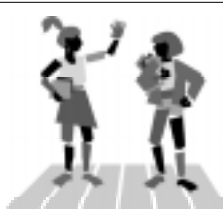
Conventions for telling stories, something children in U.S. schools are asked to do as early as kindergarten, vary tremendously among different cultural groups (Heath, 1983; Michaels, 1981). Some children use a “topic-associating” or episodic style, in which a story is composed of a string of personal anecdotes. The relevance of the anecdotes to the topic may seem obscure because there is no evident beginning, middle, or end as commonly thought of. This style is typical of many African American students (Heath, 1983).

When working with some African American children, especially from urban areas, majority culture teachers can have trouble seeing the point(s) a child is making and predicting where a story is going. Some teachers are

more accepting of children who use a storytelling approach known as *topic-centered* rather than the *topic-associating* style. A topic-centered approach to writing a story establishes a primary topic and structures the story around it.

African American teachers can be helpful to non-African American teachers in understanding and interpreting topic-associating children's accounts, because even excellent non-African American teachers have difficulties with this approach to storytelling (see also Ball, 1997). They do not perceive it as having a pattern because the pattern is so different.

The Ball research illustrates the value of having teachers from different backgrounds get together to evaluate student work and to talk not only about the common standards they hold, but about how to interpret



#### DID YOU KNOW?

Researcher Arnetha Ball had four European American teachers and four African American teachers, all enrolled in the same graduate program of a highly regarded private university, evaluate the writing of six fifth-/sixth-grade students. Three of the students were African American, two were native English-speaking Hispanic students, and one was European American. All students were considered academic achievers, with one African American student and the European American student described as “high achievers.” The teachers were trained to use a holistic rubric, as well as scales evaluating different aspects of writing. On average, European American teachers gave higher scores to the European American student than to the other students, while the scores of the African American teachers did not differ by ethnic group, in general. However, the African American teachers were more impressed with the texts written by the high-achieving African American student. *All* the teachers tended to evaluate essays written in what Ball describes as “standard academic organizational patterns” more positively.

Ball notes that the African American and Hispanic students structured their writing not only according to standard academic patterns, but also at times according to other patterns common in the oral language of the groups they came from. It is likely that the European American teachers did not recognize these patterns as legitimate or desirable (or even see them as patterns, perhaps), while the African American teachers did, which could account for why they didn't penalize students for using them. Nevertheless, it appears that all eight teachers had an “ideal” in mind, represented by the standard academic organizational pattern.

(Ball, 1997)

differences in performance that originate in cultural and linguistic differences. As the Ball example demonstrates, a recognition of differences doesn't have to mean a lowering of standards; rather, it may call for eliminating unnecessary penalties in judging students' work.

As is clear from this example, it is not only English language learners whose writing may be misjudged by some teachers, particularly when teachers from one cultural background are scoring the writing of students from another cultural background. There are cultural differences in both oral and written discourse patterns. Many decades ago, Bartlett (1932) studied the narrative patterns of some American Indian groups and was surprised to discover that a good story was structured in "fours." There were four episodes, or four opportunities for a protagonist to solve a problem, or four deeds a hero needed to accomplish. Just think of how different this is from the accepted rhythm of the typical Western European-based fairy tale (with three bears and three chairs and beds, or three pigs, or three attempts to solve the usual problem). Not only that, but the beginning, middle, and end of American Indian narratives were not readily evident to someone outside the culture. To the uninitiated, the stories seemed formless. We become so conditioned to hearing a familiar pattern over and over again that we find exceptions to it strange: a two-episode story sounds incomplete; a four-episode story sounds too long. There is, of course, nothing inherently superior to a story with a particular number of episodes or even a "clear" beginning, middle, and end.

Some literary traditions focus much more on setting than on plot, others heavily on character. But imagine a middle school student's grade on a story-writing assignment if he spent 50 percent of his space setting the scene or 75 percent on character development. A teacher unfamiliar with any of these traditions can only judge writing based on them as deviant from the norms she does know. Acknowledging the validity of other writing or speaking conventions does not preclude teaching additional conventions. Students do need to learn the discourse patterns of the mainstream U.S. culture.

---

**Table 4: Language Patterns**  
 (Adapted from Bennett and Bennett  
 of the Intercultural Communication Institute, 1993)

---

**LINEAR:**

Discussion is conducted in a straight line, developing causal connections among subpoints towards an end point, stated explicitly. Low reliance on context.

*Let's get right to the point!*

**CIRCULAR (CONTEXTUAL):**

Discussion is conducted in a circular movement, developing context around the main point, which is often left unstated. High reliance on shared context.

*Once you have the whole picture you'll know what I mean!*

**DIRECT:**

Meaning is conveyed through explicit statements made directly to the people involved, with little reliance on contextual factors such as situation and timing.

*What you see is what you get!*

**INDIRECT:**

Meaning is conveyed by suggestion, implication, nonverbal behavior, and other contextual clues; for instance, statements intended for one person may be made within earshot to a different person.

*What you get is what you manage to see/hear!*

**ATTACHED:**

Issues are discussed with feeling and emotion, conveying the speaker's personal stake in the issue and the outcome.

*If it's important, it's worth getting worked up over!*

**DETACHED:**

Issues are discussed with calmness and objectivity, conveying the speaker's ability to weigh all the factors impersonally.

*If it's important, it shouldn't be tainted by personal bias!*

---

Our challenge as educators is to help students walk successfully between two or more forms and patterns of communication. We must teach them to know when to use the conventions and structures of each.

## The Dialect Issue

A *dialect* is one version of a language that differs systematically from another version of the same language. Everyone speaks a dialect, and no single dialect of English — or any other language — is superior to another, according to its complexity or ability to communicate whatever its speakers need to communicate to each other. However, some dialects are accorded greater or lesser *social* value than others according to the status of the groups that speak them. Among some of the broadly recognized dialects that tend to be devalued by members of the dominant society are the Cockney dialect in England and, in the United States, Ebonics, or Black Vernacular English, and the New York City dialect (Wolfram, 1981). Dialects vary according to pronunciation, syntax, and word meanings, as well as in features of pitch, volume, and pacing. Dialects also generally differ in conventions, such as how stories are structured and told, rules of conversation, how questions are posed and answered, and how various forms of written language are organized.

Traditionally, dialects have been viewed as mutually intelligible versions of the same language. In other words, despite variations in pronunciation or usage, a speaker of Southern English can generally understand a speaker from the Northwest. Numerous people speak more than one dialect, choosing the most appropriate one based on the situation or the person with whom they are speaking. In fact, most of us speak differently at home or in casual conversation with friends than we do at work or in more formal situations — moving from more formal to less formal speech. But for some people, these differences are greater because they speak two distinctly different dialects.

The home language of many African American children offers an important example of the complexity of dialects and of how a dialect can operate much like a foreign language for many native-English-speaking students. Although teachers cannot be expected to be linguists, they do need to have some basic understanding of their students' dialects so they can make sense of students' oral and written communication.

Known as “Black English Vernacular,” “Ebonics,” “African American Vernacular English,” or “Black Language” (Perry & Delpit, 1998; Fillmore, 1997), this dialect or language (scholars argue over the designation) is like all other linguistic systems in that it has rules and is capable of serving all of the intellectual and social needs of its speakers. Black Language has multiple forms — oral and written, formal and informal, vernacular and literary (Perry, 1998). Its form and use have been influenced by West African languages, as well as by the social circumstances surrounding African Americans' histories in the United States. For example, deletion of the final consonant in a consonant

cluster (wes' for west or col' for cold) brings English words more in line with the form of words in some West African languages (Smith, 1998). That said, Black Language speakers can certainly learn to hear these final consonants and spell words in "school" ways.

The oratorical styles of rhythm, rhyme, metaphor, and repetition used by African American preachers are some of the other distinctive elements of Black Language (Perry, 1998). Teachers can capitalize on these aspects of language in their instruction by encouraging students to use these elements in their creative writing assignments (poetry, essays, or narrative). They can provide opportunities for classroom presentations in which students explicitly use rhetorical strategies from their home dialect (e.g., a persuasive speech). In assessment, recognition of rhetorical strategies that come from what students have been exposed to in their communities adds to fairness in grading and scoring. "Standard" patterns are not always required, particularly when effective communication within a certain context is the goal.

Black Language is distinguished by many other conventions as well, including particular structures for storytelling or narrative writing (Ball, 1997; Michaels & Cazden, 1986; Heath, 1983). Although teachers cannot be expected to be linguists, they do need to have some basic understanding of their students' dialects, so they can make sense of students' oral and written communication. If a student uses an unfamiliar narrative structure in telling a story, a teacher may immediately assume that the student is using a defective "mainstream" approach. But the student may actually be using a fully developed strategy associated with his or her home dialect or language. Similarly, when a student uses grammar or spelling that reflects the dialect he or she has learned, these differences need to be understood not as errors that indicate learning problems, but as alternative forms that are correct within a different system. Students need to learn the counterparts of these forms in the new dialect, which in the case of school is what's known as "standard academic English."

If students don't use the kind of language expected in school, they need to learn how; but that does not mean they must give up their first language or dialect. Rather, they can become proficient in two dialects. In addition, we want to stress the parallels between bidialectism (proficiency with two dialects) and bilingualism (proficiency with two languages). In fact, those African American children — and others — who do not speak the "mainstream" English dialect face much the same challenge as children learning a completely different language, even though in the case of African Americans, their dialect and the "mainstream" dialect are largely mutually intelligible.



## VIGNETTE: A STORY FROM OAKLAND

In an interview printed in “Rethinking Schools” (see Section VI), Oakland, California, fifth-grade teacher Carrie Secret describes how her approach to teaching English has shifted to one in which the language and culture of her African American students are valued equally with standard academic English and the mainstream culture that dominates in most schools.

*“Our mission ... continues to be: embrace and respect Ebonics, the home language of many of our students, and use strategies that will move them to a competency level in English.*

*She describes the change: “In the past, I used the ‘fix-something-that-was-wrong’ approach. I was always calling for the children to say something correct or to fix something to make it right. I now approach the same task from a different perspective that has a more positive affect on my children.*

*“Some days I simply announce: ‘While you are working, I will be listening to how well you use English. In your groups you must call for translation if a member of your group uses an Ebonic Structure.’ Some days I say, ‘Girls, you are attending Spelman, and, boys, you are attending Morehouse College.... Today you use the language the professors use and expect you to use in your classes, and that language is English.’*

*In the midst of teaching the language that provides access to the world outside their home community, Secret also makes sure students value their home language.*

*“We read literature that has Ebonics language patterns in it. For example, last year in fifth grade we read Joyce Hansen’s *Yellow Bird and Me*, and in fourth grade we read her book *The Gift Giver*. The language was Ebonic in structure. The language was the bonding agent for students. The book just felt good to them.*

*“When writing, the students are aware that finished pieces are written in English. The use of Ebonic structures appears in many of their first drafts. When this happens I simply say, ‘You used Ebonics here. I need you to translate this thought into English.’ This kind of statement does not negate the child’s thought or language.”*



### THINGS TO CONSIDER

There are many different ways to communicate, and, as educators, we must value them all, while, at the same time, help students know how and when to use the language of the dominant culture. Which communication patterns are you most comfortable with? Do you know the styles common in the cultures of your students? How do you decide when it is acceptable to use “nonstandard” language and when “standard English” is required in your own classroom?

Lisa Delpit, an expert on culture and dialect, reminds us that standard English is the language of power and business. When we do not equip students with the skills they need to speak and write standard English, we place them at a disadvantage. Carrie Secret demonstrates how we can both recognize and respect students’ language strengths *and* learn what students need in order to master the “standard” dialect.

---

## Gender, Language, and Assessment

Language style and use are also associated with gender differences. A quick look at the differences in the experiences of boys and girls in the same classrooms provides immediate evidence that all students do not have equal opportunities to take an active part in whole-class and small-group work. When girls are recognized, they rarely get to say more than a sentence or two before a boy interrupts and takes over the conversation. Many girls also put themselves down by beginning with a disclaimer when answering a question in class. For example, they may say, “I don’t know if this is right, but...” Further, teachers respond differently to boys’ and girls’ requests for help, being more apt to coach boys to get the answer themselves while giving girls the answer directly.

Likewise, boys are more likely to be instructed how to perform tasks while girls often have tasks done for them, and boys tend to receive feedback related to a task, the content, or thought process. Girls are more likely to get feedback based on the appearance of their work. Girls sometimes get graded up because they communicate better (Hughey, 1984; Hughey & Harper, 1983). Boys may get graded down on the basis of handwriting (Sprouse & Webb, 1994; Sweedler-Brown, 1993) or, as mentioned earlier, even behavior (Bennett, Gottesman, Rock, & Cerullo, 1993).



## THINGS TO TRY

Such recognized biases in schools are just a part of the sets of experiences that tend to limit the expectations and perceptions of each gender. Some programs are addressing the needs of girls more directly. The “Teaching SMART” program in Rapid City, South Dakota, has a model program for helping girls in mathematics and science (Vobejda & Perlstein, 1998). Some insights from the program include techniques that more fully engage girls in learning. Try some of these techniques in your classrooms:

- Wait five seconds after asking a question before calling on a student to answer — teachers are likely to see more girls raise their hands. The boys’ hands typically shoot up immediately, but gradually, this practice assures a better balance. Perhaps some boys who don’t routinely respond will also start to do so.
- Don’t always let a group decide who will lead and who should record (write on the chalkboard, keep minutes, etc.). Experience shows girls will almost always end up recording because they have nicer handwriting.
- Call upon all students whether or not their hands are up.
- Help students develop alternative ways of communication in which both boys and girls are given the same opportunity to talk and be listened to.

**What did you notice when you tried some of these techniques? Did the communication patterns in your classroom change? If so, how? What are the implications for assessment?**

---

## Understanding Students' Language Proficiency: Focus on Second-Language Learners

Whatever else might affect students' communications patterns, for students who are not native English speakers, there is also the question of language proficiency. Can they understand and use English well enough to be accurately assessed in English on their understanding of content? Other aspects of communications notwithstanding, the integral role of language itself in teaching and learning makes it imperative that teachers have a solid understanding of their students' English language skills. Consider what can happen when, instead of relying on assessment data, teachers rely on their assumptions about students' language.



### VIGNETTE: FALSE ASSUMPTIONS

The following example describes the experience of an education consultant who works with teachers and students in schools. In his own words, he describes a surprising experience he had in one school in which erroneous assumptions about students' language altered a teacher's expectations of students.

*"I was on time. This was my chance to try out with students the electricity assessment that I had been developing for the last two months. Ms. Gallagher came to me, smiling:*

*'I'm glad that you're from Mexico,' Ms. Gallagher said. 'I'm concerned about two Hispanic students I have. They don't understand a word of English and I'm afraid they're getting bored. Would you mind working with them? I don't have tests in Spanish. Maybe you can tell me how they're doing in science.'*

*"When I try out my assessments in schools, I usually have students work in pairs anyway. From observing their interaction, I can tell how they interpret the directions provided and what kinds of thinking and problem-solving strategies they use in doing the assessment. No, I did not have any problem working with them.*

*“As I walked with the kids down the hallway to the corner of the school’s gym where I had set up all my materials, I tried to melt the ice. I introduced myself and told them where I grew up — in Spanish, of course. Then I asked them to tell me where they were born. After exchanging glances, they tried to respond in Spanish, but they couldn’t. I could tell they were uncomfortable that they were not fluent in Spanish. We switched to English.*

*“As they engaged in the science task they interacted all the time in English, and they had no problem reading and writing in English. In fact, their English wasn’t any better or worse than that of any other kid in the same school.*

*‘This is fun,’ said one of them as the bell rang. So engaged were they connecting batteries and bulbs that they didn’t mind missing recess.*

*“As I drove back to my lab, I wondered how many opportunities to learn had been denied to these kids because their teacher had erroneously assumed they only spoke Spanish. Then I had a scary thought: These students might be tested in the future in Spanish, under the assumption that they do not understand a word of English.”*



## THINGS TO CONSIDER

How can we account for this nearly unbelievable situation? For starters, instructional and assessment practices are often a reflection of teachers’ perceptions of other individuals and other cultural groups. Perhaps the students did not voluntarily speak up in class, and the teacher, assuming they spoke no English and thinking she was being helpful, never called on them because she didn’t want to put them “on the spot.” The students might not have even realized that the teacher thought they spoke no English. And had they known, they might still have done nothing to correct her misperception — especially if their families had come from rural areas in Mexico where, as is true in many traditional communities, students and parents do not

question the decision of a teacher, who is considered powerful and authoritative. Possibly, too, the parents, themselves, were not proficient in English and would have found it difficult to speak with the teacher.

Whatever the reasons for this misunderstanding, it would have been avoided entirely with appropriate and timely language assessment. Determining a student's proficiency in English should be a first step in deciding whether to gauge his or her learning with an English-based assessment. If a student isn't fully proficient, an assessment in English cannot be expected to yield valid results. Unfortunately, determining if a student has greater proficiency in English or in another language isn't as easy as one might think.

---

## The Challenge of Assessing English Language Skills

The necessity for all students to become fully proficient in “standard” English is clear, as is the need to assess their progress in doing so. Because demonstrating what one has learned is heavily dependent on language proficiency, it's critical that teachers be able to accurately judge students' proficiency and the degree to which it either supports or impedes students' ability to show what they know in the academic realm.

One study of how to accurately determine English proficiency found that different measures of proficiency produced a different pattern of classification. In other words, students were classified at different levels of English proficiency depending on the source of information used (Solano-Flores, Ruiz-Primo, Baxter, Othman, & Shavelson, 1994). On one kind of test, a student may score as “fully English proficient,” while on another, the same student may score as “limited English proficient.” Clearly, then, it's important to use multiple sources of information in assessing language proficiency. In addition to a formal assessment, teachers may well want to draw on more informal assessments, such as observations in both social and academic situations and parent reports of language use at home.

There are different types of proficiency, and Olmedo (1981) notes that when testing members of linguistic minority groups on academic learning, “it is essential to understand the kind and degree of bilingualism of the individual being tested.” Conversational fluency is not

the same as academic fluency. On the playground, an English language learner might chatter away in English with his or her friends, but that doesn't mean the child's English language skills would be adequate for an academic test. In a similar vein, a native Spanish speaker, for example, could score well on one assessment of his native language skills, but if the assessment didn't address *academic language* proficiency, the results could *falsely* predict greater success with Spanish than is warranted. Academic language proficiency is heavily dependent on having used a language for academic purposes. When children cease learning academics in their first language at school, that language may not continue to develop as a tool for academic learning — unless opportunities outside of school support such language learning.

If determinations about English proficiency are based on insufficient or superficial information, a student may be improperly identified as English proficient. If so, student performance would be assessed without taking into account any language limitations, and the student's achievement would *look* worse than if the student had been assessed in his or her first language. Conversely, without a thorough assessment, a student who is actually proficient in English could be improperly identified as an English language learner, akin to the experience of the two children above whose teacher thought they spoke no English. One result could be that teacher expectations about the student's performance would be lowered, and the student might not be given the opportunity to adequately demonstrate his or her learning.

## Distinguishing Between Language Proficiency and Academic Proficiency

The following vignette demonstrates some of the issues to consider in academic assessment of English language learners.



### VIGNETTE: THE HISTORY ASSESSMENT

*Mr. Strauss’s fourth grade has just finished a unit on California history. He wants to assess his group of largely immigrant Latino students on how well they have understood the role of the Spanish missionaries in colonizing California. He intends to capitalize on their interests and make the assessment motivating, so he creates an open-ended question with room for some choice in response. He cues the students to some elements of knowledge that he does expect to see in their writing. He gives them 50 minutes, so that they can think about what they want to write first. Here is his assessment prompt:*

*“There were many players in California history who had a lasting influence on our state. Among these were the Spanish missionaries, particularly Father Junipero Serra. Write an essay of at least two pages on the Spanish missionaries, explaining what they did and how it affected California and the people who lived there at the time. You might want to imagine that you were alive in those days. Tell what you think are the pros and cons of the actions of these missionaries. What was good, and what was bad? How have their actions affected the California of today?”*

*Mr. Strauss knows that for most of his immigrant Latino students the actions of the Spanish missionaries are very personally meaningful and positive (even though he has had them do research on the effects of the missionaries’ actions on the California Indians). After all, these missionaries brought Christianity and literacy to the New World. He wants to assess what they have learned and encourage them to consider all the historical evidence. He hopes he will elicit different*

*points of view, and then students can discuss aloud why they took the perspective they did. But this is the first time he has given an extended writing assessment. Usually he gives quizzes that require short answers. He isn't prepared for the limited output he gets from his students. Most of them produce only a few sentences, and those are not connected well to address the questions in the prompt. This is their first year in an English-only classroom, but they all did well on their language proficiency test, and he knows they were very involved in this particular unit.*



### THINGS TO CONSIDER

What is going on? Mr. Strauss has matched his assessment to what he has taught. He has fairly clear ideas of his learning targets, and he believes he has chosen an assessment method that will allow students to show what they know. He has considered how to make the assessment tap students' points of view and interests. Why have they done so poorly? After all, this particular history should be easy for them to connect with on a personal level. These are some of the possibilities:

- Mr. Strauss's assessment relies very heavily on written language proficiency. Students who have been transitioned to English-only instruction may be proficient in oral language but lag behind in writing skills (in part because of instructional practices).
- On-demand writing tasks — where students don't have time to make outlines and drafts but go almost directly to a final product — are often difficult for students required to write in their second language. Process writing, which has been widely adopted and may be what students are used to, provides for planning and multiple drafts and usually includes peer support. Writing an essay on a complex topic in 50 minutes is an entirely different task.
- Mr. Strauss has nicely given a context to his prompt. He didn't just launch into a direct question. However, in order to comprehend the prompt, the student has to

read some fairly complicated sentences and grapple with at least four questions. Students reading in a second language are likely to take longer to get an accurate sense of the overall problem and formulate a plan for responding to it than native English speakers.

- When students write in a second language, they may have to pay undue attention to the mechanics of writing — something that takes attention away from monitoring the overall product (how well they are communicating or addressing what they meant to, how well their argument is stacking up, etc.).

---

Assessments designed for second language learners can go wrong in so many ways. Sometimes the directions on a test are ambiguous or require understanding of complex syntax. This is troublesome because, if a student does not frame a problem correctly from the outset, his or her solution is likely to be flawed (Durán, 1985). In the case of Mr. Strauss and his students, one of the problems may have been that the assessment prompt, or set of questions, was just too linguistically complicated. Students had to wade through a lot to get the full picture of what Mr. Strauss expected them to do.

But even if they understood the prompt and had a lot of knowledge about the topic, creating a formal essay would be very demanding for students who have just been transitioned to English language arts. The reality is that even students who pass language proficiency tests at high levels are still developing as users of English, and they cannot be expected to perform like native English speakers. Mr. Strauss could have scaffolded the assessment task by having students brainstorm in small groups, make an outline of the main points they wanted to address, and use a graphic organizer (just a simple T-chart, like Ms. Altcheck used) to list pros and cons before writing the actual essay.

Sometimes teachers choose oral assessment for the reasons outlined above. But oral assessments are not always easier than written ones. Instructions given orally can put too great a demand on memory. When giving instructions or posing a question orally to students who are still mastering English, there are several ways teachers can adjust the language demands. They can slow their speech, give context cues, and relate what they are saying to students' past experience (Watson, Northcutt, & Rydell, 1989).

Teachers need knowledge of how language factors may cloud the picture of academic achievement and how to create assessments that minimize this confusion. They need skills in mediating — or adjusting — the

administration of assessments so performances will reflect the true learning of students. Another very simple step Mr. Strauss could have taken would have been to walk his students through his assessment prompt and have some of them paraphrase what they thought they were asked to do.

## Responding to Student Writing Errors

Much assessment *is* done through writing, and it's natural to notice student errors in writing (whether writing is the object of assessment or the vehicle for assessing learning in another subject area). But not all errors merit the attention they receive (Leki, 1992; Valdés, 1991). When it comes to evaluating student performance, teachers need skill in determining which kinds of student errors indicate learning problems and which are simply normal developmental blips along the way to language mastery.

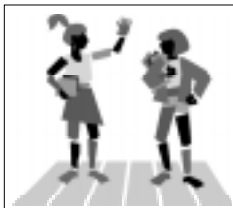
There need to be established criteria for what should count as *serious* errors. For instance, even though omission of articles (as in, "My family had picnic this weekend.") grates on the ears of many English teachers, they need to realize that this is a relatively trivial error. In addition, when teachers are familiar with students' first languages, they can look at student work with an understanding of how the first language is "coming through," versus focusing unduly on errors (Sweedler-Brown, 1993). In the case of articles, for instance, many Asian languages do not use them at all. Even English language learners who have progressed to "advanced" status as English speakers and writers may continue to struggle with these issues. General education teachers can benefit from working closely with those who teach English as a second language or bilingual teachers. Both groups speak and/or know a great deal about the languages of their students.

The same benefits can be derived from collaboration between African American teachers familiar with Black Language and their non-African American colleagues. Students who exhibit features of Black Language in their writing are also learning a second code, and similar understanding needs to be brought to bear by teachers.

## Assessment Adaptation and Translation Issues

When a student performs poorly on an academic assessment, we often do not know if the performance reflects poor learning or simply an inadequate mastery of language (García & Pearson, 1994; Hamayan & Damico, 1991). A student may have met requirements for English-only instruction but still take longer to process ideas in English than in his or her first language. For this reason, a timed test or time-limited assessment of any kind may penalize an English language learner. Vocabulary in a second language may be less developed than that of a native speaker. A person may know a word such as “credit” or “bark” but not know all the common meanings or the connotations of the word.

Students who speak English as a second language may perform differently on an assessment, depending on which language they are allowed (or choose) to use. In an investigation on the use of monolingual (English-only or Spanish-only) and bilingual (English and Spanish)



### DID YOU KNOW?

It has been suggested that language minority students' schooling aptitude is underestimated if testing is conducted only in English. On the other hand, when an assessment is translated from one language to another, its validity is seriously jeopardized. Its level of difficulty may change, or it may be assessing something entirely different from the original from which it was translated (Olmedo, 1981).

response formats in science performance assessments, Solano-Flores, Ruiz-Primo, Baxter, and Shavelson (1992) found evidence suggesting that not only test format but also the language chosen for responses are factors in student performance.

In another investigation, Solano-Flores, Ruiz-Primo, Baxter, Othman, and Shavelson (1994) attempted to obtain more accurate information on language proficiency. They examined three types of data on English proficiency: standardized English proficiency scores provided by schools, patterns of English use in different sociocultural contexts, and patterns of use in bilingual

test formats (for example, reading the items in one language and responding in another). They found that each measure of English proficiency produced a different pattern of classification. In other words, the students were classified at different levels of English proficiency

depending on the source of information used (Solano-Flores, Ruiz-Primo, Baxter, Othman, & Shavelson, 1994).

The main implication of these results is that any decision about a student's English proficiency should be made only after different sources of information have been carefully examined. If decisions about student English proficiency are based on insufficient or superficial information, a student may be improperly identified as English proficient. Her performance will be assessed without taking into account her language limitations, so her achievement will look worse than if she were assessed in her first language. Conversely, an English-proficient student may be improperly identified as an English language learner. Expectations about her performance will be lowered, and she will not be given the opportunity to adequately demonstrate her learning. Even "accepted" approaches to assessing English language learners can give inaccurate results unless the students' sociocultural context is taken into account.

In this era of testing, assessments are being modified for use with students for whom the instruments weren't originally created. Among these modifications are test accommodations, such as changes on administration procedure or administration time, which may be made for English language learners or students with certain disabilities. There are also test adaptations, such as format changes or translation. The quality and validity of the scores produced when accommodations are used are still under investigation.

One of the most common adaptations for English language learners is translation. In fact, many assume that the whole issue of language proficiency in assessment can be avoided simply by translating tests into the language in which a student is most proficient. But that's a risky assumption. Studies in which different language versions of the same assessment are used show that even a single word whose translation is slightly inaccurate may produce unintended interpretations of an item and alter its psychometric properties, giving an advantage to students tested in the assessment's original language.

There are also some sociocultural reasons for why even a perfectly translated assessment may, nonetheless, yield inaccurate results. In one study of bilingual assessment formats, researchers created a Spanish version of assessments developed in English by using a standard translation process commonly accepted as valid (Solano-Flores, Ruiz-Primo, Baxter, & Shavelson, 1992). In this process, an experienced translator translated the original assessments into Spanish; a panel of bilingual scholars reviewed the Spanish versions and translated them back into English to monitor retention of the original meaning; and, finally, the Spanish versions were refined based on a series of try-outs with a sample of Spanish-speaking students.

From the scholarly point of view, the translation was accurate and the grammar impeccable. But some Latino English language learners were unfamiliar with some Spanish words that the researchers had *assumed* were part of the students' everyday language. Because the assessments had originally been developed in English, the sociocultural background of the English language learners hadn't been considered during the development process. The scores of these English language learners would have been considered dependable because the standard translation procedure had been used, but, in reality, the assessments would not have been valid because, for these particular students, the assessment inadvertently tested students' Spanish vocabulary knowledge rather than their academic content knowledge.

This is one more reminder that when the sociocultural background of minority students is not properly considered, even commendable efforts intended to promote equitable testing may yield flawed results from assessments of dubious validity.

Because a translated version of an assessment does not undergo the same process of development and refinement as the original, the overall quality of the translated version may be inferior. Some researchers are investigating the limitations and possibilities of a model for concurrently developing content area assessments in English and Spanish. The skills and knowledge to be assessed are specified, with the content and format for each assessment task agreed upon. Then, when assessment developers make a change in one language version of the task, (e.g., English), they make an equivalent change in the other language, (e.g., Spanish). That way, at the end of the process the two sets of assessment tasks are comparable across languages (Solano-Flores, Trumbull, & Nelson-Barber, 2000).



#### VIGNETTE: EXAMPLES OF ADAPTATION ERRORS

What follow are examples of the wide variety of adaptation-related errors that have been found in tests, with an explanation of the circumstances that led to these particular errors. All these examples come from tests that have been translated from English for Spanish-speaking English language learners. But, as you'll see, such errors could easily arise in translations for other languages as well.

- *A science prompt for second-graders shows an illustration of a crab and asks the student to indicate the crab’s pincers. The prompt is translated as: “Señala la panza del cangrejo.” The translator meant to write “pinza” (pincer), but instead wrote “panza,” which means “tummy” or “belly.” The inadvertent substitution of one vowel significantly changed the meaning of the prompt.*
- *A reading diagnostic test in English, considered appropriate for first-graders, is translated into Spanish to diagnose reading skills of Spanish-speaking English learners. The English word “bat,” which is monosyllabic and has three letters, has to be translated as “murciélagó,” which has four syllables and 10 letters. Although the translation is correct, any conclusions that can be made about native Spanish speakers based on this translation would be objectionable because of the differences in the reading levels.*
- *A science test that has been available in English for several years is translated to serve native Spanish speakers. To ensure comparability across languages, the format and appearance (e.g., number of pages, location of items and illustrations) of the Spanish version must be the same as the format and appearance of the English version. Because printed Spanish takes about 25 percent more characters than printed English, the only way the translator finds to meet these requirements is to reduce the font size of the text in the Spanish version. Although the two language versions have similar appearances, the Spanish readers may have more difficulty reading the test because of the small font size.*
- *A mathematics test that required one year to develop is about to be printed. The publisher is then asked to produce a Spanish version that will be given to English language learners during the same administration period. The company hires only one translator, who is given one week to deliver the Spanish version. With such a tight timeline, the wording of the prompts cannot be reviewed and refined as carefully as it was for the English version. As a result of the rush, several errors are made. For example, in a series of word*

*problems in the Spanish version, the last sentence of the text is missing. Yet two of the five multiple-choice items still refer to that sentence. The mistake is only noticed and corrected long after students have been assessed and placement decisions made based on the assessment results.*



### THINGS TO CONSIDER

Determining the best way to assess English language learners, even those who have been judged English-proficient, is difficult. Teachers may want to offer students choices in the languages of the assessment wherever possible.

No equitable testing can be attained if the first languages of minority students do not receive the same treatment as English throughout the assessment development process. Obviously, translation and other adaptation procedures need to be improved.

---

## Summary

It is clear that we have many issues to consider in order to ensure fair and useful assessment of our students' learning. Understanding students' culture-based differences in ways of using language, students' language proficiency, and how language is involved in the whole assessment process are keys to valid assessment. For some students, it may take years to make sense of and be comfortable participating in the kinds of interchanges considered routine in American classrooms. By that time, many of them will have been lost to the system — or, at the very least, misjudged. Other methods need to be cultivated with such students as early in their schooling as possible to equip them with the skills that will enable them to navigate successfully between their own culture and that of school.

Assessments can reveal what students know and can do at a given point, or they can actually obscure students' learning. In particular, when a student produces nothing or very little on a test, it is difficult to know what is going on. Does he or she really not have the skills or knowledge necessary to perform well on the test? Or is the assessment not doing its job?

Unfortunately, it is often easier to blame the student than the assessment, even though we need to look to both for explanations about low performance. Even if the problem lies with the student, distinguishing among causes can be difficult. Less than optimal learning can be caused by the following student factors (among other things):

- A mismatch between student language needs and language of instruction
- Lack of effort
- Lack of ability
- Lack of motivation
- A learning disability
- Conscious or unconscious resistance (when a student's individual social identity seems to be threatened by what he or she is required to do to participate in the classroom)

The assessment factors we have identified are those that interact with student factors. It is, of course, possible to have an assessment that is invalid on other grounds: the developer may not have selected a set of questions or tasks that adequately sample the domain being assessed, and wrong conclusions may be made about a student's competence in the domain (let's say mathematical reasoning). Or the assessment may

be too difficult for nearly all students, on the basis of what has been taught. Here, we are concerned with assessment factors that can lead to misjudgments about students' opportunity to learn or student differences from "mainstream" expectations that have nothing to do with the skills being assessed.

When teachers are knowledgeable about how language, culture, and a host of other factors interact with assessment, they can take steps (and support their districts to take steps) to improve the ways assessments are designed, administered, scored, and interpreted.