4

Fairness

The goal of assessment is to obtain a sound inference about what a student knows, understands, and can do. Such inferences can then lead to good decisions that enhance student learning and can also inform teachers, the public, and policy makers. While the principles of reliability and validity have served us well as the foundation of sound assessment, it has recently become evident that issues of fairness are equally important.

There are many ways to think about fairness in relation to educational assessment, and it is not possible to discuss all these complex issues in this book. A full consideration would explore testing in relation to broader social goals, and legal and ethical issues. I will be concerned with fairness as reflected in the nature of assessments and the use of information generated from assessments in classrooms and schools.

WHAT IS FAIRNESS?

The term *fairness* can be defined from several perspectives (Heubert & Hauser, 1999). It can be defined broadly as a condition or situation in which assessments are not unduly influenced by factors unrelated to the learning objectives or standards that are being measured. Fairness can also be described as equitable and just treatment of those being assessed. The term *bias* is often thought of as the opposite of fairness. Obviously, assessments should be free from bias, whether that bias is based on gender, race, socioeconomic status (SES), or other characteristics that may influence the performance being assessed. If some students have an advantage because of factors unrelated to what is being assessed, then the assessment is not fair.

Fair assessments are unbiased and nondiscriminatory, uninfluenced by irrelevant factors such as gender or race and by subjective factors such as the bias of a scorer. In other words, in a fair assessment, students have the opportunity to demonstrate their learning in a way such that their performance is not distorted by their race, gender, ethnic background, disability, or other factors unrelated to the purpose of the assessment. A fair assessment is also characterized by scoring that is not affected by these factors.

Unfair assessments result in performances that both underestimate and overestimate the traits being measured. A common example of unfair assessment is when particular groups of examinees are put at a disadvantage. This frequently occurs when there is something about the content of the assessment that makes it more difficult for students with certain characteristics. For example, suppose a test is designed to assess students' writing skills by asking them to respond to the following prompt:

> Write a short story about a boy who practices very hard to be good in basketball and makes the team.

This type of prompt may be easier for boys to respond to than for girls. Girls may score lower than boys not because they don't have the writing skills but because they are less familiar with this sport. This suggests that the prompt is biased against girls. On the other hand, boys may score higher because they have experience to draw from in writing their short stories. In either case, there is some distortion in the performance caused by a factor unrelated to what is being assessed. Another good example is a reading comprehension assessment in which the content of the reading passage favors one group over another. If the content is about a sailing experience, then examinees who live in a community in which sailing is ubiquitous would probably have an easier time reading and understanding the passage than would examinees who know little about sailing.

Although it is impossible to completely remove all unfair aspects of an assessment for every student, teachers and administrators can do much to ensure that assessments are as fair as possible. We can also understand the various ways fairness can affect our interpretations and uses of assessment results. This is particularly important given the increasingly diverse characteristics of students and increased emphasis on identification of students with special needs. Because of these conditions, fairness is just as important as the standbys of validity and reliability.

FAIRNESS RELATED TO STUDENT KNOWLEDGE OF LEARNING TARGETS AND ASSESSMENTS

Have you ever taken a test and then thought, "If I had known the test was going to cover this content, I would have studied it"? I know I have. A fair assessment

is one in which it is clear to students what will and will not be tested, what test format will be used, and how the test will be scored. The goal is not to trick, fool, or outguess students about what is assessed. Rather, teachers need to be clear and relatively specific about the nature of the target students need to learn and how that performance will be scored. When this is communicated to all students *prior to instruction* and in a *public way*, students, as well as parents, know what needs to be learned. This approach helps students know what to study and focus on as well as how to study. For example, if students know that they will need to show their work on a mathematics test, they will practice showing their work and, it is hoped, receive feedback about what they show. If students know that they will have an essay test that will be graded with a known set of scoring criteria, they will pull together their thoughts in a way that organizes the content to respond to the criteria, rather than simply memorizing content.

Let's look at another example in more depth to illustrate the importance of specificity and clarity in learning targets, the test itself, and scoring criteria. Suppose you are going to teach a unit on cultural diversity. Here is the first learning target you identify:

Students will be able to identify and learn about different cultures.

You tell students that they will be tested on this learning target at the end of the unit. This information, although shared with students, is so general and vague that it indicates little about what will be learned, the test format, and the scoring criteria. Students are left guessing about what they really need to learn, what will be on the test, how they will be tested, and how the test will be scored. Consider the next more specific target and description of assessment:

Students will identify characteristics of several cultures and be able to show how they are both similar and different.

You inform students that the assessment will be a constructed-response test in which students will list the characteristics of each of the cultures and indicate ways they are similar and ways they are different. With this information, students have a much better idea about what to learn and how their learning will be assessed. There is much less guessing based on idiosyncratic student differences, which makes the assessment fairer.

Here, the target is even more specific:

Students will identify six characteristics of three cultures, explain how each culture differs from the others, and explain the implications of these differences and similarities on freedom of expression and tolerance.

Now students have a clear idea what they will learn and how they will be required to demonstrate their learning. This type of target is fairer because it takes the ambiguity out of the process, replacing it with clarity and specificity.

Another good illustration of the importance of the targets is learning to drive. Students take a driver education course with the expectation that certain specific skills will be tested. It would be unfair if students were taught to drive using an automatic transmission and then, when asked to demonstrate their skills, had to do so with a manual transmission. What *would* be fair is telling students at the beginning of the course that they had to learn to drive competently with both a manual and an automatic transmission.

One of the positive consequences of ensuring that students know the learning targets, the test format, and scoring criteria in advance is that it can help motivate students and help them obtain a learning goal orientation. Recent research has indicated that students are intrinsically motivated when they learn for mastery, rather than for mere performance (Ames, 1992; Dweck & Leggett, 1988). With a *learning* or *mastery* goal orientation, students are motivated by a focus on mastering a task according to established standards, developing new skills, improving competence, and gaining understanding and insight. In addition, when students assume a mastery orientation, they are more likely to see a significant link between their effort, feedback, and the outcome. This link promotes more internal attributions for success (e.g., reasons for success such as ability and effort). Finally, mastery-oriented goals lead to greater effort and involvement as well as greater interest and positive attitudes (Pintrich & Schunk, 1996).

In contrast, when students have vague information about the learning targets, the type of assessment, and scoring criteria, they tend to assume a performance goal orientation. Students are likely to focus on the consequences of the outcome, that is, a grade or some other recognition or reward, with relatively little concern about the level of understanding or learning. When only general information about the targets and assessments is provided, students find it more difficult to see the link between their effort, feedback, and the outcome.

The principle of goal orientation is nicely illustrated with gymnastics. The targets are clear and specific, and the criteria by which gymnasts are judged are public and well understood by the athletes and coaches. This results in a learning orientation that encourages intrinsic motivation. When gymnasts are judged, the assessment is usually fair because the athletes know that the same rules apply to all and they know, in advance, how their performances will be judged. Gymnasts prepare a specific routine to demonstrate their skills, knowing ahead of time that the routine is appropriate. Can you imagine gymnasts attending a meet not knowing the nature of the specific routine they should perform or how it will be judged?

Figure 4.1 shows how student knowledge of specific learning targets and the nature of the assessments used to measure the knowledge lead to more positive student motivation, engagement, attitudes, and achievement.

FAIRNESS RELATED TO OPPORTUNITY TO LEARN

Opportunity to learn is a phrase that means students have had adequate time and appropriate instruction to enable them to obtain proficiency or mastery. It

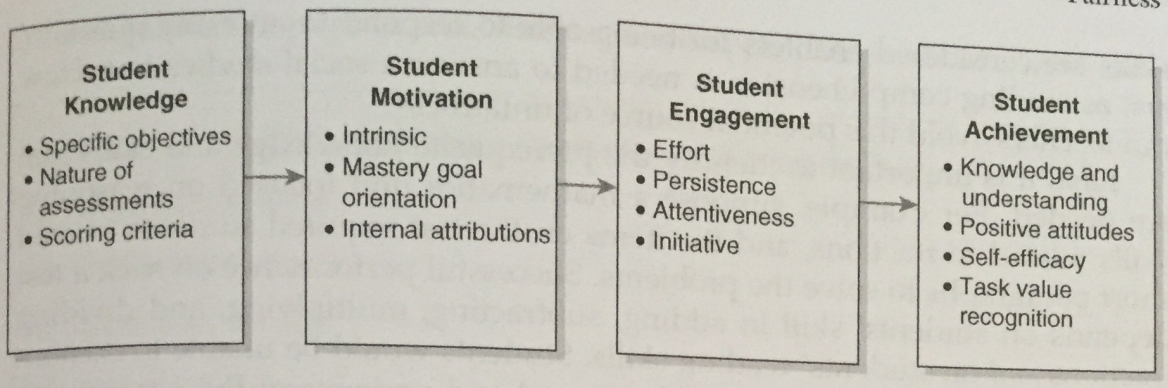

Figure 4.1 Positive Effects of Students' Knowledge About Targets and Assessments

simply is not fair if students are held accountable for knowledge and skills that have not been taught or that have not been part of the curriculum. This is especially troublesome when interpretations about the quality of schools or teaching are made on the basis of student achievement tests. Obviously, when students have not had an adequate opportunity to learn the knowledge and skills covered by the test, they are likely to get low scores. Serious decisions based on these low scores, such as school accreditation, teacher employment, and withholding of high school diplomas, are invalid due to lack of opportunity to learn. It is not so much a matter of accuracy because students indeed may not have the knowledge and skills, but the consequences are inappropriate.

At the classroom level, opportunity to learn is directly related to what teachers do and the instructional resources provided. Did students know *how* to learn the targeted knowledge and skills? Were there adequate resources for all students? Was there adequate supervision at home for homework? Did the lessons implemented cover what was on the test? Has illness or absenteeism made it difficult for some students to learn the material? How adequate was the instruction? Is it fair to hold students accountable for poor instruction? These types of questions raise issues related to opportunity to learn. Obviously, opportunity to learn is a matter of judgment and of degree, but there should at least be guidelines to identify situations that clearly limit students' opportunities so that this information can become part of the interpretation of the scores. For example, it would not make much sense to make a negative judgment about a teacher or school on the basis of student test scores based on knowledge and skills that the teacher or school simply did not include as part of instruction.

FAIRNESS RELATED TO STUDENT PREREQUISITE KNOWLEDGE AND SKILLS

It is not fair to assess students on content or skills that require prerequisite knowledge, understanding, or skills that they do not possess. Such knowledge and skills are often referred to as *enabling behaviors* because they are necessary but not sufficient for demonstrating the targeted learning. For instance, writing

skills are considered enablers for being able to respond to an essay question, just as reading comprehension is needed to answer a social studies test. How can teachers avoid this potential source of unfairness?

First, it is important to identify the prerequisite knowledge and skills that are needed. For example, suppose a mathematics unit focused on reasoning skills related to fractions, and the items on the test required students to read short paragraphs to solve the problems. Successful performance on such a test depends on students' skill in adding, subtracting, multiplying, and dividing fractions *and* on students' reading skills. Students would be unable to demonstrate their reasoning if they could not read and understand the paragraphs, even if they were proficient at working with fractions.

Second, teachers need to have a good understanding of the prerequisite knowledge and skills that their students can demonstrate. Sometimes, this is done with formal pretests or other structured assessment, but more often, teachers use informal assignments and oral questioning to get a sense of whether students have the needed knowledge and skills.

A different but equally important type of prerequisite skill is concerned with test taking. *Test-taking* skills are those that allow students to maximize their performance by not being distracted by format or approach. For example, if a new question format is used, students need to become familiar with that format prior to the assessment. If students are going to be using Scantron forms for the first time, they need practice in using these forms. More general test-taking skills include the following:

- Paying careful attention to general directions and how answers are to be made
- Paying attention when reading or listening to items
- Pacing so that not too much time is spent on one item
- Being willing to skip difficult items initially and return to them later
- Answering all items if all are included in the scoring
- Learning how to guess the correct answer
- Learning how to omit wrong answers on multiple-choice tests rather than looking for the right answer
- Checking so that item and answer numbers match
- Checking answers if time permits
- Organizing essay answers before writing them
- Realizing that some items will be very difficult and not being too concerned when this occurs
- Knowing the scoring criteria for performance assessments
- Knowing acceptable formats for completing performance assessments
- Knowing how to prepare for the test
- Knowing how to handle test anxiety

Another set of skills has been identified as *test-wise* skills. These skills help the examinee identify correct answers by errors in test questions or by clues to the correct answer. Such deficiencies are common in poorly constructed multiple-choice test items. Should students be taught the following skills to maximize scores?

- Looking for grammatical clues such as inappropriate use of "a" and "an."

In a study of the effect of training on performance, the training is an

a. dependent variable

b. independent variable

c. continuous variable

d. control variable

- Looking for vague words such as "often" or "usually" that may indicate the correct answer. Absolutes are typically incorrect.

In an experiment, the independent variable is

a. measured at the end

b. always matched with a control variable

c. almost always categorical

d. the most important variable for understanding the result

- Looking for options that are longest or most precise.

If a quasi-experimental study examines the effect of training on performance, what possible extraneous variable would need to be controlled if at all possible?

a. Differences between the groups that could account for differences obtained on the dependent variable

b. The situations in which the performance is tested

c. The directions given to the subjects

d. The time of day of the training

Should students be trained in these types of skills? Although all students clearly need test-taking skills, there may be some hesitancy in teaching students to be test wise. What is important for fairness is that all students have the same degree of test-wise skills. To ensure this, it seems to me that it is desirable to inform all students about these types of skills. More generally, students should be familiar with the format and type of question and response that will be needed on an assessment. This is often accomplished by giving students practice questions or showing them examples done by students in previous years. This doesn't mean teaching the test, that is, using examples during test preparation

that are identical to what will be on the test. But it does make sense to teach *to* the test in the sense of teaching students the content and skills that will be assessed.

FAIRNESS RELATED TO ABSENCE OF BIAS IN ASSESSMENT TASKS AND SCORING

A fair assessment is one in which neither the assessment tasks nor the scoring contains bias. This type of influence is perhaps the best known source of unfairness in testing. Bias can be defined as qualities of the assessment that distort performance because of the student's ethnicity, gender, race, religious background, SES, or other characteristics. Such distortion, as pointed out earlier, can distort by either raising or lowering test scores, but usually bias is associated with a negative impact (hence the phrase "biased against"). Popham (2007) has identified two forms of assessment bias: offensiveness and unfair penalization.

Offensiveness

Offensiveness occurs when the content of the assessment offends, upsets, angers, distresses, or otherwise creates negative emotions for students of particular subgroups. The negative effect influences the performance of these students, lowering their scores and reducing validity. The distressed students are distracted from what is being assessed and focus more on the offensive content. Offensiveness is particularly unfair when stereotypes of specific subgroups are present. Suppose a test question portrays women in low-status, low-paying jobs and men in professional positions. Women taking the test may be offended by the negative portrayal. The distress leads to less than optimal performance, resulting in scores that underestimate the actual knowledge of the students. Some men taking the test will also be offended by such content. Table 4.1 shows some additional examples of test items that may create offensiveness.

In large-scale testing, item writing and review procedures are used to eliminate any offensiveness in the content of the assessments. In classroom testing, however, it is more likely that such offensiveness will occur. Teachers simply do not have the time or resources to conduct systematic reviews of test content. For assessments that have important consequences, it is advisable for teachers to ask a colleague to review the content to look specifically for offensiveness, as well as other types of bias. Because teachers are often unaware that the content may be biased by unconsciously including wording or characterizations that may offend some students, a review by a colleague is helpful.

Unfair Penalization

Unfair penalization refers to bias due to the content of the assessment. This makes the assessment more difficult for some students than for others. In other

Table 4.1 Examples of Offensiveness in the Content of Assessments

Ethnicity	Juan picks beans for a living. He receives 20 cents for every bushel he picks. Juan picked 20 bushels a day for 2 weeks. How much should he be paid for his work?
Gender	The president of General Mix has held his present position for 10 years. His secretary has been with him for all 10 years. She has received a 4 percent raise each year, whereas the president has received a 10 percent raise each year. If the president started at a salary of $100,000 and his secretary at $10,000, what is the current difference in their salaries?
Race	Write a 2-page essay on the following: African American teenagers constitute 10 percent of the population, but 30 percent of the crime in this city is committed by African American teenagers. How do you explain the difference in these percentages?
Religion	Write a 200-word essay on how right-wing Republican Christians have influenced the outcome of the past two presidential elections with their extremist views.

words, bias is present when a disadvantage is given to one group or individual because of gender, SES, race, language, or other characteristic. This is the type of unfairness when the simple phrase "this is a biased test" is used. The content makes it harder for some students to score well because of factors unrelated to what they are learning in school.

Suppose you take a test that measures your aptitude by using mostly rural, farm-oriented content. The items cover such topics as types of cows and pigs, farm equipment, and winter crops. If you grew up in a city or suburb, do you think your score would be as high as the scores of students who grew up on a farm? Similarly, is it fair to compare students whose primary language is Spanish with students whose primary language is English on English oral reading skills? Do test items containing sports content give boys an unfair advantage over girls because the boys are more familiar with sports? In each case, membership in a particular group or background unrelated to instruction influences the score. Further examples of unfair penalization are illustrated in Table 4.2.

Suppose test scores are different for specific subgroups. This may or may not mean that the assessment is biased. For example, if Hispanic students score lower, overall, on the SAT, this does not mean that these tests are culturally biased and unfairly penalize Hispanic students. The actual content of the tests needs to be analyzed to determine bias. While publishing companies are careful to exclude any content that may unfairly penalize students of certain groups, it is virtually impossible to remove all types of bias from such tests. Questions need to be written with some type of content, and invariably this content will unfairly penalize some students to some degree.

In the increasingly diverse culture of the United States, student differences reflected in vocabulary, prior experience, skills, and values may influence both

Table 4.2 Examples of Unfair Penalization

Socioeconomic background	Students are required to pass a computer competency test to be promoted to high school.
Socioeconomic background	Students are required to work in small groups to plan a trip from their school to Washington, DC. In their plans, they must include expenses and an itinerary of activities. The plans will be graded on comprehensiveness and practicality.
Religion	What is the Koran?
Location	Write a persuasive essay about using hiking as recreation. In your answer, compare hiking with sailing.
Gender	Dale Earnhardt's car traveled at 180.5 miles per hour for the entire race. How many seconds did it take him to finish the 200-mile race?

formal and informal assessments. Consider the impact of the following cultural differences (McMillan, 2007a):

- Knowledge from the immediate environment of the student provides a vocabulary as well as an indication of the importance or relevance of assessment tasks (e.g., large city, ethnic neighborhood, rural, coastal, and farm).
- Depending on the culture in which the student lives, there may be different norms and rules for sharing beliefs, discussion, asking questions, taking turns, and expressing opinions.
- Respect and politeness may be expressed differently by students from different cultures (e.g., silence, squinting as a way to say no, looking up or down when asked a question, and not looking into another's eyes when answering a question).
- Learning style differences may influence a student's confidence and motivation to complete certain assessment tasks (e.g., preferences for working alone or in a group, learning by listening or reading, ability to think analytically or globally, and tendency to answer reflectively or impulsively).

The influences of these differences will be minimized to the extent that teachers and administrators first understand them, then review assessments for possible bias, and finally incorporate possible limitations because of bias in their interpretations of student performance. The differential impact of cultures is also minimized by using multiple assessments with varying formats. This helps students show their progress toward achieving the learning target and results in more valid inferences about their progress. If one assessment technique or approach advantages students from one type of background, another technique may be a disadvantage to those same students. Using different types of assessments provides a balance to the others. For example, students who perform poorly on an oral test may perform well on a written test. This points out an important principle in all assessments: *Never rely solely on one method of assessment*. This doesn't mean that you should arbitrarily select different assessment methods. Use a variety of assessments that provide the fairest indication of performance for all students.

FAIRNESS RELATED TO AVOIDING STEREOTYPES

In making judgments about students, it is only natural to form and use beliefs related to how students are likely to perform. These beliefs about what students are capable of knowing or doing are called *expectations*. Expectations are not necessarily bad. Realistic, accurate expectations are helpful in designing appropriate instruction. Expectations that are biased because of membership in a particular group, however, need to be avoided. When such expectations occur, the stereotypes will influence expectations and the nature of subjective judgments. Here are some examples of stereotypes in education:

"Jocks aren't very bright."

"Girls do better in most subjects than boys."

"Girls do poorly in mathematics."

"Kids from the south side are great athletes but are weak academically."

"A single-parent home means the father won't be involved."

In its most negative form, inappropriate use of stereotypes is self-fulfilling and detrimental. For example, if a teacher has a stereotype about students from a certain ethnic background and that stereotype is translated into behaviors toward the students, then the students may interpret that behavior as an indicator of their capability. Should students perform consistently with the communicated level of performance, the teacher or administrator will in turn interpret this as evidence to reinforce the original stereotype. A useful way to think about such stereotypes is to consider when they are made in relation to assessments, that is, prior to, during, or after instruction.

Stereotypes Prior to Instruction or Assessment

A hallmark of an effective teacher is to match instructional activities with the capabilities that students bring to school. Teachers need to assess these capabilities prior to making final decisions about instruction. Such assessment occurs before school begins, continues during the first week or two of the new school year, and occurs again during the year when beginning new units. During these "preinstructional" times, information is gathered and interpreted to answer such questions as the following:

Do students have the content knowledge and skills to handle the new material?

In what aspects of the content will students be most interested?

How can I take into account backgrounds of students to maximize motivation?

How much heterogeneity will there be in the class? How can I accommodate students who are behind others?

Stereotypes develop when interpretations and conclusions are based on fragmented or partial information. Consider the following examples:

"John comes from a single-parent family; he will have difficulty completing his homework."

"Tanya is from the low-income area of the city; she will be behind in mathematics."

"This class has mostly boys; it's going to be difficult to control."

"These Hispanic students from Mott Middle School will need remediation."

"Carol has dirty clothes; she probably didn't get a good breakfast and will probably have difficulty concentrating in class."

In each case, judgments are made on the basis of stereotypes, rather than on evidence specific to each individual or group.

To design the most effective instructional experiences, it is helpful to gather as much evidence as possible about student backgrounds and capabilities. When this is done systematically by reviewing school records, test scores, and recommendations from previous teachers, it is more likely that undesirable stereotypes will be avoided. Stereotypes are also avoided as teachers interact with and get to know each student. This interaction occurs during the first week or so and, when added to other information, can provide a complete and accurate assessment.

Stereotypes During Formative Assessment

As instructional activities are being implemented, there is a need to constantly gather "evidence" from students that is focused on how much students are paying attention and learning. This continuous monitoring provides feedback to the teacher to assess progress toward understanding the content or accomplishing the skill. Teachers typically use informal observation of students to accomplish formative assessments. This involves looking at and listening to students and then interpreting their behavior. Both verbal and nonverbal student behavior is important. The goal is to obtain an accurate picture of where students are in relation to the lesson and learning targets, and stereotypes will interfere with an accurate interpretation of these observations.

Stereotypes can influence formative assessment in a general, continuous way or can be more specific to particular situations. If teachers have a stereotype about boys being on task less than girls, in general, teachers may tend to monitor boys more closely than girls. If there is an expectation that students who live in poverty will, in general, have more difficulty learning the content, then that stereotype may affect the way the teacher interprets questions asked by these students. Asian American students are often stereotyped as high achievers. If this means that silence from these students is interpreted to mean that they understand, while silence from other students means they are bored, then the stereotype has interfered with an accurate formative assessment.

On a more specific level, teachers can form stereotypes about certain students. General opinions of individual students as "able" or "bright" or "smart," on the one hand, may contrast with opinions of other students labeled as "unable," "not very bright," or even "dumb." This stereotyping can form a "halo" for each student that may distort subsequent evaluations to be consistent with the nature of the halo. For example, a student viewed as bright may give an answer that will be interpreted on the basis of the halo as much as on the basis of the quality of the answer. It is also possible for teachers' formative assessments to be influenced by primacy effects, in which initial impressions have a distorting effect on later assessments, and by recency effects, in which interpretations are unduly influenced by the most recent observation.

Stereotypes During Summative Assessment

Stereotypes distort summative assessments most in the scoring of student responses. When constructed-response assessments are scored, teachers use their judgments of the responses to interpret student progress and learning. These judgments require subjectivity to a certain extent, and it is in this subjectivity that stereotypes can lead to unfair evaluations. Suppose a teacher is grading responses to an essay question. Will knowing the name of the student completing the response mean that some type of stereotype of this student will influence the evaluation (e.g., "She never does well on these types of assignments" or "She always does great on essays")? This is definitely possible. Is it likely that a teacher would evaluate a performance assessment more positively for some students than for others solely on the basis of characteristics unrelated to the assessment? For example, will student athletes receive lower evaluations because a teacher has a stereotype that they are more concerned about their sport than about an upcoming test? Are students who belong to the drama club stereotyped as better actors and actresses, so that when judgments are made about who will be assigned different parts in a play, this stereotype influences the judges to select the drama club students for the best roles?

These types of stereotypes can be avoided if scoring criteria are specific and explicit. The more general and vague the scoring criteria, the more opportunity there is for stereotypes to influence the judgments that are made. When grading students' constructed-responses, it is best to grade one question at a time for all students, and, if possible, grade them without knowing the names of the students. Examples or outlines of answers should be generated to act as anchors in the judgment process.

It is also helpful to review the performance of test takers of different races, gender, and ethnic background to determine if performance is related to these characteristics. If, for example, all the female examinees do well and all the male examinees do poorly, there may be something in the scoring related to gender stereotypes that is unfair. As indicated earlier, differential performance by different groups doesn't mean there is bias, but if there is bias, it will be reflected in this way. The summative assessments, then, provide data that may indicate bias or stereotypes.

FAIRNESS AS IT RELATES TO ACCOMMODATING SPECIAL NEEDS

One of the most challenging aspects of teaching is accommodating students with mild to moderate disabilities who are now routinely mainstreamed into regular classes. From the standpoint of assessment of these students, it is important for legal, ethical, and instructional reasons to adapt assessment practices so that they are fair and unbiased. Legally, teachers are responsible for gathering information, through assessment, to identify students who may become eligible for special education services. Assessments are used by teachers to provide information needed to determine if students are making satisfactory progress toward meeting learning targets specified in their Individual Education Programs (IEPs). For both of these responsibilities, the law requires that the selection and administration of assessments must not be racially or culturally discriminatory. At a minimum, the law requires the following:

1. Personnel administering tests must be trained.
2. Assessments must be in the child's native language.
3. Assessments must identify specific needs, not a single, general indication of ability or aptitude.
4. Assessment materials and administration must accurately indicate aptitude or achievement without discriminating against the child's disability.
5. No single score or procedure is sufficient as a sole criterion for determining an IEP.
6. A multidisciplinary team needs to assess the child in all areas related to the suspected disability.

These provisions mean that assessment must be planned and implemented so that the score is determined by the trait being assessed and not by the disability. That is, it would be unfair to use a test written in English to determine that a student, whose primary language is Chinese, has low ability or aptitude. It would also be unfair to conclude that a student with a fine motor disability did not demonstrate understanding as reflected in an essay question because there was insufficient time for the student to write the answer. In other words, it is illegal for the score of the trait being assessed to be influenced by the disability.

Impacts of Disabilities on Assessment

Beyond these legal requirements, teachers need to make appropriate accommodations in classroom assessments because many disabilities affect test-taking abilities. Without such accommodations, scores are unlikely to be valid or reliable. Students with disabilities have specific difficulties that are directly related to assessment. These factors are summarized in Table 4.3.

Table 4.3 Factors That Affect the Assessment of Students With Disabilities

Factor	*Impact on Assessment*
Poor comprehension	Understanding directions; completing assessments
Poor auditory skills	Understanding oral directions, assessment tasks, and questions; being distracted by noises
Poor visual skills	Understanding written directions, assessment tasks, and questions; decoding symbols and letters; being distracted by visual cues
Time constraints	Finishing assessments
Anxiety	Finishing assessments and being able to think clearly; demonstrating best work
Embarrassment	Finishing assessments; being reluctant to ask questions
Variability of behavior	Finishing assessments; demonstrating best work

Many students with even mild disabilities have difficulty with comprehension. This means that they may not understand directions well or remember a sequence of steps required to complete a task. This is especially troublesome for abstract tasks that require reasoning or other thinking skills. For example, a question such as "How is the Canadian government different from a socialistic form of government?" would be much more difficult than "What are the characteristics of a socialist form of government?" Auditory and visual difficulties can exacerbate limits to comprehension. If students have trouble processing verbal information quickly and accurately, or if visual and perceptual disabilities make it hard to discriminate letters and symbols, then comprehension is affected. During tests, many students with disabilities will be easily distracted by visual cues such as gestures or motions of others, which disrupt their visual focus and concentration.

Time can be a major difficulty for many students with disabilities. It simply takes these students longer to complete tests because of limitations related to how quickly they can understand and process information and record their answers. Timed or speeded tests, for example, may lead to increased levels of anxiety, especially if students are concerned that their disability will make it difficult to complete the task within given time limits.

Students with disabilities may be more sensitive to feelings of embarrassment than are other students. To avoid embarrassment, they often want to hide or disguise their disability so that they will not be singled out by peers. As a result, they may want to be treated like all other students. They may not ask questions about directions they do not understand or may hand in a test when others do, even if they have not finished. Finally, the behavior of students with disabilities may vary significantly from one setting to another or from one time to another. Consequently, teachers need to be flexible with assessments and realize that at any one time, a disability may pose extreme difficulties for the student.

Assessment Accommodations

Assessment accommodations for students with disabilities can be grouped into three categories: test directions, test construction, and test administration.

Directions

Directions can be modified in the following ways:

- Read written directions slowly, and give students ample opportunity to ask questions.
- Keep directions simple and short.
- Give examples of how to answer questions if students are not familiar with the format.
- Give separate directions for each section of the test.
- Give one direction in each sentence.
- Check students' understanding of the directions.
- Monitor the students during the assessment.

Test Construction

Tests should be constructed to include plenty of white space, a font size that is not too small, and double spacing. Different sections should be clearly differentiated, with only one type of question on each page. Each page of the test should be numbered. Other accommodations to the format of the test depend on the type of item, as illustrated in the following examples.

Short-Answer, Essay, and Completion Items. Short-answer and essay items may be difficult for students with disabilities because of the organization, reasoning, and writing skills required. Long, complicated questions should be avoided. If words such as "compare," "contrast," and "discuss" are used, they need to be clearly defined. Limit the number of questions, and allow students to outline their answers. Some students may need to give an audio-recorded, rather than a written, answer, and all students with disabilities will need sufficient time to complete their answers. Word banks can be provided on a separate sheet to aid memory. Provide plenty of space for students to record their answers.

Multiple-Choice and Binary-Choice Items. Instruct students to circle their answers, rather than writing the letter next to the item or transferring their answers to a separate sheet. Make sure that response categories are arranged vertically, not horizontally. Usually, the number of alternatives in a multiple-choice item should be limited to four, and wording such as "a and b but not d" and "either a or c" should be avoided. Any negatively stated stems or alternatives should be used sparingly. If words such as "not" are used, they should be highlighted with bold print and/or underlines.

Performance Assessments. The first accommodation with performance assessments that will need to be made is with the directions. What is expected needs to be clearly explained with examples and a reasonable time frame. Certain skills may need to be modified if the disability interferes directly with performance of the skill. Assistance needs to be provided in cases in which the disability makes it difficult for the student to perform the skill. For example, a speech impediment may affect a student's ability to give a speech or oral report. In this situation, the student may need assistance in organizing and delivering the speech or report.

Portfolios. Portfolios may be an ideal type of assessment for students with disabilities because the assignments and requirements can be individualized to show progress in whatever time frame is appropriate. Portfolios could be adapted by modifying requirements to include those that would be least affected by the disability. Reflections by both the student and teacher could be included that specifically address progress despite the presence of the disability.

Adaptations in Administration of Assessments

Many of the adaptations that need to be made in the administration of an assessment are dependent on the nature of the disability. In general, the goal is to use procedures that lessen or remove the negative impact of the disability on the trait being assessed. Some of the suggestions for adaptations are based on common sense (e.g., for students who have difficulty hearing, be sure that the directions are written; for students with a visual difficulty, make sure directions are given orally). It is best to place a "Testing: Do Not Disturb" sign on the classroom door to discourage visitors and other distractions. Some students may need to be removed from the regular classroom and taken to a separate room in which distractions are minimal. If someone can monitor the assessment, the student will have more opportunities to ask questions and be less likely to be embarrassed when asking for clarification or other assistance.

Classroom teachers who are unsure about appropriate assessment accommodations should check with the special education specialist in the school. This will help the teachers understand the nature of the disabilities and what specific accommodations are appropriate.

In Table 4.4, suggested adaptations are summarized according to different types of disabilities. For many students, however, several of these difficulties will need to be addressed (e.g., students with an auditory disability may also have comprehension, anxiety, and time constraint difficulties).

OVERALL JUDGMENTS ABOUT FAIRNESS