

ASSESSING STUDENT LEARNING IN MODEL-BASED INSTRUCTION

As you know by now, quality physical education programs are based on standards of learning or other outcomes determined by a school's teaching staff. But standards alone are not enough. Teachers must be able to document that students are meeting the standards to a satisfactory level. In order to provide that documentation, teachers will need to assess student learning on a regular basis. Teachers also should use assessment to inform their instruction. If the assessment indicates that some or all of the students failed to meet a lesson's objectives, a teacher should modify her instruction accordingly. This chapter will describe the many things to consider in making good assessments in physical education within a model-based instruction approach. Assessment in physical education is done for three important reasons:

1. To describe how much learning has taken place within a given amount of instructional time (a lesson, a unit, a year, or the entire program) and to use that information to plan future lessons and units.
2. To judge or evaluate the quality of that learning (usually for grading).



3. To make decisions for improving instruction when one or more standards are not being met to a satisfactory level.

ALIGNING ASSESSMENT WITH STANDARDS

Assessment provides the crucial link between standards or other defined outcomes for a program and the program's success. As you recall from Chapter 1, the NASPE standards for physical education (NASPE, 2004) are widely accepted as the best and most comprehensive list of the key learning outcomes for contemporary physical education programs. Because standards express "what should be learned," they also provide the best starting point for conducting assessment to answer the next key question: "How can teachers document that the standards have been learned?" If a school, district, or entire state has adopted the NASPE standards, for example, the standards should guide the assessment process from start to finish. Because of the wide acceptance of the NASPE standards, these standards will guide this chapter. When you begin your teaching career, your school or district may have adopted a different set of standards. However, this chapter's example of the NASPE standards will help you understand how standards can be used to guide and assess instruction. This chapter will also give you an abundant set of resources and examples to draw on beyond this book. Many of those resources will be listed at the end of the chapter.

KEY ASSESSMENT CONCEPTS AND TERMS

The first step in learning about assessment in physical education is to become familiar with the various concepts and terms used in that process.

Assessment and Evaluation

Two terms are commonly used to describe the process of documenting what students have learned. Be aware that the terms do have different meanings, but they are often used interchangeably, and this can be confusing. *Assessment* includes those processes used to gather and analyze information to document what students have learned in a lesson, unit, or program. Assessment describes or measures what has been learned and uses that information to plan for future lessons and units. The next—but different—step, *evaluation*, involves a determination of the worth or value of that information. That evaluation is done by the teacher, a student, or some other person who can judge "how good" the learning was. Let me give you a common example to show the difference between these concepts. You are a player on a softball team, and your team scores five runs in the game. "Five runs" is an assessment to *describe* your total for the game. However, it does not tell you if scoring five runs was good enough to win. To evaluate *how well* your team met its goal of winning the game, you must know the other team's score. If it is more than five, your

team did
well enou
Gradi
step inclu
many poi
happens
indicating
(i.e., "D,
So, i
mance, c
uses that
knowled
tion. In r
most tea
better do
takes pl
class pa
assessm
of the b

Timing

Assess
of inst
7.1 for
inform

1. Co
the un
be use
on the
assess
the te
One r

Exan

CONTI

Self-
each

Chec
task

Cher
seve

team did not play well enough to win; if it is less than five, your team played well enough to win.

Grading starts with assessment and leads to evaluation. The assessment step includes determining how many correct answers a student gave, how many points she earned, or what percentage score she got. The evaluation step happens when the teacher converts those answers, points, or scores to letters indicating whether a student passed or failed and, if he passed, to what degree (i.e., "D," "C," "B," or "A").

So, if a teacher gathers information about students' knowledge, performance, or behavior related to a standard, that is **assessment**. If the teacher uses that information to make a judgment about the worth or quality of that knowledge, performance, or behavior, he has taken the next step of **evaluation**. In reality, most of what is done in physical education today is assessment; most teachers want to know what their students have learned in order to make better decisions for their instruction and their programs. What little evaluation takes place usually measures non-performance criteria such as dressing out, class participation, effort, and behavior. To reduce the potential for confusion, *assessment* will be used as the key operating term in this chapter and the rest of the book.

Timing of Assessments

Assessment information can be gathered according to three time plans in a unit of instruction: (1) continuous, (2) formative, and (3) summative. See Exhibit 7.1 for examples. Each time plan will give the teacher different amounts of information and determine what can be done with that information.

1. *Continuous assessment* occurs automatically during each learning task in the unit. It provides the teacher with "up-to-the-minute" information that can be used to make decisions about task progression—whether to keep students on the present task longer or advance them to the next task. Continuous assessment produces smaller amounts of information, but the timeliness allows the teacher to make changes for the next lesson or even within the current one. One model in Part Two, Personalized System for Instruction (PSI), uses only

Examples of three time plans for assessment.

EXHIBIT

7.1

CONTINUOUS	FORMATIVE	SUMMATIVE
Self-check performance tasks each day	Short weekly quizzes	Written final exam
Checking for understanding after task presentation	Partner-checked tasks after each skill component	Skills test at the end of the unit
Checking for target heart rate several times in class	Weekly log of fitness activity	Fitness testing at the end of the unit

continuous assessment techniques to provide information about every student in every lesson.

2. *Formative assessments* are conducted periodically during the unit of instruction, perhaps every second or third lesson. They provide the teacher with “midstream” feedback on student learning, so that changes can be made during the unit—before it is too late. Formative assessments can be of short or medium duration, giving the teacher enough information to make changes without forcing him to sacrifice instructional time.

3. *Summative assessments* are conducted at the end of the unit and allow the teacher to determine how much learning occurred during the entire time of instruction. They usually generate larger amounts of information because they cover most or all of the intended learning outcomes. Summative assessments typically require more student preparation time (e.g., studying for tests) and more teacher organization time. Although summative assessments provide more information than continuous or formative techniques, they come at the end of the unit, so the teacher cannot use that information until the next time she teaches that unit.

Norm-Based and Criterion-Based Assessments

All assessments generate a score or other indicator of student performance, and all scores need some level of interpretation by the teacher to be understood. How that interpretation is made will depend on whether the assessment technique used was norm-based or criterion-based.

Norm-based assessments require the collection of a large number of scores on standardized tests, such as fitness tests or skills tests. Scores are then compiled into subgroups according to some common factors of test takers, usually age and gender (e.g., 10-year-old boys, 14-year-old girls). The large number of scores for each group will generate a *normal distribution* of scores that allows any one student's score to be compared with the scores of other students of the same age and gender. Raw scores can be determined and reported, but normed test scores are usually reported as a percentage of students who scored above and below each individual student. For example, being in the 85th percentile of 10-year-old boys means that 84 percent of all other similar test takers scored lower and 15 percent of all other similar test takers scored the same or higher. If a teacher uses a norm-based test, she can then compare her students' performance to all other similar students who have taken the test. Some normed tests can help determine if a student is developing at or near the expected maturation rate in areas like motor skills, fitness, and social development.

Criterion-based assessments compare a student's performance to standards (the criteria) deemed appropriate for mastery in a particular area. Scores from these assessments usually offer a number, or percentage, for the amount of material each student has mastered. The performance standards are determined by the person who uses the assessment technique and reflect one's professional judgment about the quality of each score. A teacher who makes

her own te
75 percent
points to g
The value
what the
outcomes

Organiza

All assess
to imple
advance
of neede

Info
little cla
lar flow
observe
when a
on the
enjoyment
called
tion stu
present
to reca
they a

For
for th
of the
ments
often
some

Exam

INFO

Che
sen

Ask
abl

Asl
he

As
bo

out every student

ring the unit of
vide the teacher
ges can be made
can be of short
to make changes

nit and allow the
ne entire time of
ion because they
utive assessments
ng for tests) and
ssments provide
they come at the
until the next time

ent performance,
her to be under-
er the assessment

number of scores
es are then com-
st takers, usually
large number of
cores that allows
er students of the
rted, but normed
ho scored above
e 85th percentile
test takers scored
e same or higher.
students' perfor-
ome normed tests
ected maturation

rmance to stan-
particular area.
rcentage, for the
nce standards are
and reflect one's
cher who makes

her own tests and sets the passing grade for a written test on volleyball rules at 75 percent, or who determines that students must score 45 out of 50 possible points to get an A on a tennis skills test, is using criterion-based assessment. The value of criterion-based tests is that they can be written to include only what the teacher wants covered in them, thus reflecting the specific kinds of outcomes most sought by the teacher.

Organizational Plans for Assessment

All assessment techniques require the teacher to make an organizational plan to implement them. Assessment doesn't just happen; it must be thought out in advance and implemented in a way that matches the amount and complexity of needed information.

Informal assessments are those that require little advance planning and little class time. Many informal assessments can actually happen in the regular flow of class in a way that is almost undetectable by students and outside observers. Of course, the teacher is well aware of the assessments and controls when and how they occur. The teacher can use many informal techniques on the spot to make decisions about task progression or to gauge student enjoyment of the current activity. One effective informal assessment strategy called *checking for understanding* is used to determine how much information students have retained during a short period of class time, such as a task presentation or explanation of game rules. The teacher simply asks students to recall what was just said or shown to them, which lets the teacher know if they are ready to go on from there.

Formal assessments require more planning and implementation time both for the teacher and the students. Students are typically given advance notice of these assessments and often extra time to prepare for them. Formal assessments include fitness tests, skills tests, or group projects. Formal assessments often count in students' grades or other performance reports. Exhibit 7.2 lists some examples of formal and informal assessments.

Examples of informal and formal assessment plans.

EXHIBIT 7.2

INFORMAL	FORMAL
Checking for understanding after a task presentation	Giving a written quiz on key elements of a skill
Asking for a show of hands of those who were able to make five shots on goal in this drill	Giving a skills test on soccer shooting, with the teacher counting makes and misses
Asking students, "How many of you feel your heart pounding hard right now?"	Directing students to take their pulse and calculate to see if they were in their target zone
Asking students, "How many of you like skateboarding?"	Asking students to write five things they like about skateboarding

TRADITIONAL ASSESSMENTS IN PHYSICAL EDUCATION

There are many techniques for assessing learning in physical education, regardless of whether they are designed to be formative or summative, norm-based or criterion-based, formal or informal. Assessments can be divided into two broad categories: traditional and alternative. Traditional assessments have been used in physical education for many years, as their label indicates. Alternative assessments have been recently developed to provide information about learning that is quite different from that provided by traditional techniques. Keep in mind that neither type of assessment is inherently better than the other—they simply produce different kinds of information about learning. It is up to the teacher to decide the best type of assessment to use in individual situations.

Traditional assessments in physical education follow three primary principles, derived from the subdiscipline of measurement and evaluation (Wood, 1996):

1. Establish (and formally state) appropriate instructional objectives.
2. Use appropriate (validated and reliable) tests to measure characteristics related to the instructional objectives.
3. Develop an evaluation (grading) scheme that reflects attainment of instructional objectives. (pp. 202–203)

This traditional assessment process involves an approach that resembles data collection and analysis for research purposes. It typically involves rigorous standards in all parts of the process, to ensure that the data can be used with confidence. As will be discussed below, this kind of rigor can make for a “double-edged sword” for teachers who wish to use these traditional assessments.

Some of the most common forms of traditional assessment used in physical education are:

1. *Informal teacher observation.* Probably the most common type of assessment in physical education occurs when the teacher observes students as they practice skills, play games, and answer questions in class. When asked how they know when their students have learned to throw a Frisbee, the majority of teachers will respond by saying that they saw or heard evidence of that learning in class. This is not truly assessment, but it does provide the teacher with some level of confidence that learning has occurred.

2. *Standardized skills tests.* Many sports in the physical education curriculum have standardized skills tests that assess student performance abilities (Strand & Wilson, 1993). Most of these involve static tests that measure a limited range of skills needed to play each sport. A static test uses a mostly artificial display of skills that determine the student’s ability to perform in nongame situations. The front-wall volley tests in racquetball (Hensley, East, & Stillwell, 1979) and any test that calls for the student to hit a ball into a fixed target (such as tennis-serving accuracy tests) are static. Standardized skills tests are

not necessarily invalid, but teachers who use them must understand the tests' limitations for assessment purposes. That is, these tests measure only a small portion of the skills needed to play a game well, and those skills are measured in a nongame situation.

3. *Fitness tests.* Our field has a long history of testing children and youth on several parameters of physical fitness. Over the years, we have developed many standardized test batteries for school programs. The most prominent tests used today are the *Fitnessgram* protocol developed at The Cooper Institute for Aerobic Research and the *President's Challenge* test from the President's Council on Physical Fitness and Sports. With proper administration, these tests can provide students with results that help them determine their level of fitness relative to other students of the same gender and age (Safrit, 1995).

4. *Written tests.* Many teachers use written tests to assess students' cognitive knowledge in physical education. Nearly all of these tests are developed by individual teachers who use them in their own programs; we do not have standardized written tests in physical education like those developed nationally for subjects such as math, science, and reading. Written tests can include questions and problems in several formats: multiple-choice, short-answer, fill-in-the-blank, matching, diagramming, and open-ended. Each type of question or problem will prompt students to demonstrate their knowledge in a different way and to provide different types of assessment information to the teacher.

Advantages of Traditional Assessments

Within certain limits, skills tests, fitness tests, and written tests can provide useful assessment information. They usually measure learning in ways that are sufficiently objective, use accepted measuring techniques and instruments, and can be carried out with strong consistency. Standardized skills and fitness tests have met rigorous standards in their development and have been field-tested on many subjects who are similar in age, gender, and ability to those for whom the test is designed.

Most skills tests and fitness tests are normative, meaning that they have been used to measure many people in different age, gender, and ability groups, providing a large database for the test. Having that database allows a teacher to compare her students' results with those from similar test takers in other schools, states, and regions across the nation.

Traditional assessments also have the advantage of including content, procedures, and scoring developed by trained experts, so the teacher does not have to take the time to design, validate, field-test, and determine how to score the assessment. Many such tests today include software and data-entry procedures that simplify the tasks of scoring and reporting results to students. In addition, these types of assessments usually produce large amounts of numeric data that can undergo statistical treatment to generate many kinds of informative analyses, ranging from simple measures of central tendency to sophisticated tests of learning effects.

Disadvantages of Traditional Assessments

On the other hand, traditional assessments present some distinct disadvantages that have limited their use in school physical education programs. Some of the issues focus on the true validity of test design, relative to typical outcomes in physical education. For instance, does the serving accuracy test for tennis (AAHPERD, 1989) measure one's ability to play the game (a typical program outcome) or one's proficiency in a limited, controlled, and static aspect of tennis skill performance? Are health-related fitness tests good predictors of one's health risks or continued involvement in physical activity in later years, or are they one-time "snapshots" of a child's present condition with little relationship to the future? Most likely, the second answer is correct in both examples.

For a large number of teachers, the disadvantages of traditional assessments center on problems of practicality when they are used in large classes and with limited class time (Wood, 2003). The rigor with which most standardized tests generate data comes with a price; many of them must be administered by the teacher to individual students, taking time away from learning itself. This fact forces the teacher into a real trade-off—taking the extra time needed for the assessment away from instructional time—which in turn may result in lower performance scores on the assessment! Because of the time and effort needed to complete them, traditional assessments often limit the types of information they provide for the teacher and students; they tend to focus on a small portion of the intended learning outcomes, leaving many unanswered questions about student achievement.

In addition, traditional assessments often require that the data be transformed in ways that are difficult for teachers and students to understand. Statistical treatment can sometimes produce information that requires special expertise to interpret and communicate to learners. The final disadvantage of traditional assessments is that they tend to focus on identifying discrepancies between what was learned and what should have been learned, with little information for teachers and students about how to improve knowledge, performance, or fitness levels. A written test results in a score, or a percentage of items answered correctly. It does not help students improve their knowledge. Fitness tests and skills tests give students normative results, letting them know how they compare with others of the same age and gender, but there is nothing in those tests that teaches a student how to improve personal fitness levels for the next test or how to increase activity levels in daily life.

ALTERNATIVE ASSESSMENTS IN PHYSICAL EDUCATION

Educational assessment has undergone much reform in recent years. Some of that reform has been driven by professionals who believe that the disadvantages of traditional assessment outweigh its advantages. A second factor has been the expanded scope and variety of new learning outcomes in physical education that cannot be assessed with traditional techniques (Lambert, 2007). This growing movement is called *alternative assessment*, to indicate that

Assessments

essments present some distinct disadvantages in school physical education programs. Some of the problems, relative to typical outcome measures, are the serving accuracy test for the ability to play the game (a typical program is limited, controlled, and static aspect of the related fitness tests good predictors of performance in physical activity in later years, a child's present condition with little relationship to the second answer is correct in both examples). The disadvantages of traditional assessments are when they are used in large classes and the rigor with which most standardized tests must be administered taking time away from learning itself. The time-off-taking the extra time needed for a final time—which in turn may result in a poor assessment! Because of the time and effort involved, assessments often limit the types of information that can be obtained from students; they tend to focus on a small portion of the material, leaving many unanswered questions.

ments often require that the data be
t for teachers and students to under
es produce information that requires s
nicate to learners. The final disadvan
y tend to focus on identifying discrep
what should have been learned, with
ents about how to improve knowledge
ten test results in a score, or a percent
not help students improve their know
udents normative results, letting them
the same age and gender, but there is n
nt how to improve personal fitness lev
ctivity levels in daily life.

ENTS

N

ndergone much reform in recent years, ren by professionals who believe that t assessment outweigh its advantages. A e and variety of new learning outco ssessed with traditional techniques (La called *alternative assessment*, to indica

6. *Oral examinations.* Students receive questions that require a high level of knowledge and are expected to give and explain their answers. Example questions could be: "What are the advantages and disadvantages of playing zone defenses in football?" or "Why is it important to drink fluids before, during, and after exercise?"

7. *"Show and tell" presentations.* Students do some research or practice outside of class and then present what they have learned to the class.

8. *Interviews.* Students interview other students, family members, or community members about what physical education was like when they were in school.

9. *Teacher, peer, and self-observations with performance checklists.* A set of correct performance cues is taught to students and listed on a checklist. The teacher or a classmate observes a student practicing and "checks off" those parts that were performed correctly. Or a student can be recorded while practicing and then watch the video to check herself.

10. *Portfolios.* Students collect and organize various kinds of artifacts (photographs, videos, drawings, newspaper articles, and the like) that demonstrate their knowledge of a certain topic or concept (e.g., the Olympics, Title IX). The process of identifying an organizing theme for the portfolio and gathering artifacts becomes just as important as the final contents of the portfolio itself.

Rubrics

The cornerstone of alternative assessment is the scoring *rubric*, which is used to determine the quality of the learning demonstrated by students (Lund, 2000). Rubrics inform students of the performance criteria on the assessment ahead of time and serve as the basis of the teacher's review of the completed work. The rubric translates the quality of the demonstrated learning into a descriptor (e.g., "novice," "intermediate," "advanced," or "incomplete," "almost complete," "complete") or a numeric/value rating (e.g., 1/poor, 2/fair, 3/good, 4/excellent) based on criteria stated in advance.

Goodrich (1996–97) highlights seven steps in the process of creating good rubrics:

1. *Look at models.* Show students examples of good and not-so-good work from previous assessments.
2. *List and discuss criteria.* Use the models to help students make a list of what counts as quality work.
3. *Articulate gradations of quality.* Present and discuss what separates each level of quality from the others.
4. *Practice on models.* Have the students use the rubric on the models.
5. *Use self- and peer assessments.* Allow for periodic assessments as the task is being completed. Provide students with formative feedback.

6. *Revise.* Allow students time to make revisions based on the feedback given in step 5.
7. *Use teacher assessment.* Apply the rubric in the same way that students have learned it.

Exhibits 7.3–7.5 present three examples of alternative assessment tasks in physical education and the scoring rubric for each one.

Example of alternative assessment and rubric for Newcomb rules lesson.

EXHIBIT 7.3

Grade levels: 3–4 (at the end of a 4-week Newcomb unit)

Knowledge assessed: Newcomb rules

Assessment task: Three teams are made at each court. Two of those teams will play, while the third team will watch the action from the sideline and record the results and score after each volley. Teams rotate after every 10 volleys, regardless of the score. The teacher also keeps a score sheet for assessing students' answers when the teams change (this constitutes the rubric for assessing students' accuracy).

	POINT SCORED BY TEAM A OR B?	HOW?*	WHAT IS THE SCORE NOW?	DOES ANYONE ROTATE?***
Volley #1 (A serves)	A	5	1–0	N
Volley #2	A	3	2–0	N
Volley #3	B	1	2–1	Y
Volley #4	B	4	2–2	N
Volley #5	B	5	2–3	N
Volley #6	B	3	2–4	N
Volley #7	A	1	3–4	Y
Volley #8	A	3	4–4	N
Volley #9	B	1	4–5	Y
Volley #10	B	4	4–6	N

* How?

1. Foot fault
2. Net serve
3. Out of bounds
4. Ball touched too many players' hands
5. Other team did not return the ball over the net

** Does anyone rotate?

Y = YES

N = NO

Developed by Donna Hicks, physical education teacher at Hickory Flat Elementary School, McDonough, GA. Used with permission.

EXHIBIT**7.4**

Example of alternative assessment and scoring rubric for Olympic Oath and Spirit lesson.

Grade levels: Middle school

Knowledge assessed: The Olympic Oath and Spirit

Assessment task: Following a unit on Summer Olympic sports, make a multimedia collage that shows athletes competing and displaying the "Olympic Spirit."

Time allowed: Out of class, due in one week

Scoring rubric:

MEDAL LEVEL	ASSESSMENT CRITERIA
Gold medal	1. Your collage uses three or more types of media. 2. Your collage shows at least five different sports. 3. Your collage shows many types of sports, athletes, and five countries. 4. Your collage shows athletes displaying the Olympic spirit in many ways.
Silver medal	1. Your collage uses two types of media. 2. Your collage shows three or four different sports. 3. Your collage shows a few types of sports, athletes, and countries. 4. Your collage shows just a few examples of the Olympic spirit.
Bronze medal	1. One type of media is used. 2. One or two sports are displayed. 3. Only competition is shown.
Future Olympian	1. Collage includes non-Olympic examples. 2. Collage does not fill up the poster you selected. 3. Incorrect examples of the Olympic spirit are included.

Advantages of Alternative Assessments

Just as with traditional techniques, alternative assessments have advantages that teachers need to understand. Perhaps the greatest advantage of alternative assessments is their ability to be designed to monitor the specific learning standards or outcomes planned by the teacher. The teacher can feel more confident that the assessment targets what she really wanted her students to learn. Another strong advantage is that alternative assessments allow students to demonstrate their learning in a variety of ways—not just through the recall of information or through skill performance on standardized tests. Because we know that students are engaged in diverse kinds of learning activities, it makes sense that they should be able to demonstrate their knowledge in a way that matches how they actually gained it.

Example of alternative assessment and rubric for golf swing lesson.

EXHIBIT

7.5

Grade levels: High school**Knowledge assessed:** Ability to analyze key elements of the golf swing

Assessment task: Get a partner. One partner will take 10 swings with a "driver" while the other partner observes each trial. The observer will evaluate each key element on the checklist to reflect how well his/her partner performed that component each time. The observer's checks will be compared with those made by the teacher, who will also watch the practicing student. Switch roles after every 10 tries.

KEY ELEMENT	1	2	3	4	5	6	7	8	9	10
1. Stance and alignment										
2. Grip										
3. Back swing										
4. Point of contact										
5. Follow-through										

After each shot, score:

- 3 if element was fully correct
- 2 if element was almost fully correct
- 1 if element was incorrect

Scoring rubric for the observer (1 point for each analysis that matches the teacher's; possible 50 points):

45–50	Teaching professional
40–44	"Scratch" player
35–39	Low handicapper
25–34	High handicapper
24 or lower	Novice

By using a scoring rubric, students know the performance expectations for the assessment task ahead of time. They are then able to make choices and devise strategies for the *process* of completing the assessment as well as the final demonstration of their knowledge. In other words, the rubric fosters the ability to *learn how to learn*. Rubrics provide another advantage. They include a common set of criteria used by the teacher to assess student work. Rubrics lend a sufficient degree of objectivity to the review process and make for more consistent assessments within and among classes.

Finally, alternative assessment, particularly through the rubric, has the ability to provide students with feedback about their learning and to identify ways for them to improve performance.

Disadvantages of Alternative Assessments

Although alternative assessments provide unique, creative, and diverse ways to assess student learning in physical education, they are not without disadvantages. Because they are customized, it takes time for a teacher to design each one for her class. The examples shown earlier can give a teacher some ideas for assessment, but unless those examples are perfect matches, they will need to be adapted or completely redesigned for use by other teachers. The design stage may also require some trial-and-error time to get the assessment just right for its intended use.

Alternative assessments typically take longer for students to complete and for the teacher to review. Many of these tasks require extended time for students to plan their work, conceptualize answers, organize materials, and develop the final product. All these steps contribute to the learning process itself, but they do take time to accomplish. Similarly, the teacher will need some time to review each completed assessment—usually more time than it would take to score written tests and skills tests.

AUTHENTIC ASSESSMENT FOR PHYSICAL EDUCATION

Authentic assessment, sometimes called *performance assessment*, refers to the practice of making assessments that involve, or simulate, the real-life settings in which knowledge taught in school will be applied (Lambert, 2007). Some examples will help to clarify this concept. Health-related fitness tests (e.g., Fitnessgram, President's Council) measure student abilities on several parameters that are *indicators* of present fitness levels. They do not assess health-related *behaviors and choices* made at other times during the school day, after school, and at home that more accurately determine patterns of healthy living. If a stated goal of the physical education program is to promote active and healthy lifestyles, then an authentic assessment of that learning would occur at the times and places that students make health-related choices for their behavior. Student logs that record the foods they eat and the moderate to vigorous activity they engage in would be an appropriate authentic assessment of physical fitness behavior and choices.

If the goal is to teach students how to play the game of softball, then an authentic assessment would take place during an actual game, not with a skills test in a nongame context. In this case, an authentic assessment could be based on students' game statistics or a performance checklist completed as the game is played. The Game Performance Assessment Instrument (GPAI softball version) developed by Griffin, Mitchell, and Oslin (1997) does precisely that by analyzing players' performance in several categories of involvement, skills, strategies, and decision making during game play. The GPAI will be explained in much more detail in Chapter 14.

If the goal is to assess junior high students' dance performance ability, then an authentic assessment could be designed during their participation in a class

and diverse ways to
without disadvan-
ner to design each
teacher some ideas
es, they will need
chers. The design
e assessment just

ents to complete
extended time for
ze materials, and
learning process
teacher will need
more time than it

ssessment, refers
or simulate, the
l will be applied
concept. Health-
measure student
ent fitness levels.
de at other times
accurately deter-
ysical education
authentic assess-
at students make
record the foods
in would be an
or and choices.

softball, then an
ame, not with a
essment could be
completed as the
nt (GPAI softball
es precisely that
olvement, skills,
will be explained

unce ability, then
pation in a class

dance production at the school. Asking students to demonstrate their cognitive knowledge or their appreciation for dance with a multimedia collage would be an alternative assessment technique, but that is not nearly as authentic as performing in a production. I emphasize that authenticity is determined by the *degree* to which the assessment calls for students to demonstrate real-life applications of their knowledge. Rarely is any assessment technique totally non-authentic or entirely authentic.

Physical education teachers need to know that all assessments, traditional or alternative, and at whatever degree of authenticity, are essentially tools used to understand what students have learned from instruction in a content unit or program. The key here is to understand what each assessment tool can do and when it is appropriate to use it. Teachers should not start by saying, "I am going to use an alternative, authentic assessment in this upcoming dance unit." Rather, the teacher should follow a deductive process that starts with "What and how will students learn in this unit?" which leads to "How do I want them to demonstrate that new knowledge for me?" and then "Which assessment technique will provide the best demonstration of their knowledge?" From that, the teacher will select the right assessment tool for the right assessment job.

ASSESSMENT IN MODEL-BASED INSTRUCTION

All of the assessment concepts and techniques described so far in this chapter have been presented as declarative knowledge, key things that all teachers should know *about* assessment. Effective teachers who use instructional models will need to have that kind of knowledge as well as good procedural and conditional knowledge in order to select and carry out the best assessment strategies within the model used for a unit. Exhibit 7.6 shows some assessment strategies for each of the models in Part Two of this book, along with suggested times to implement those strategies. These assessment techniques will be explained in more detail in the assessment section of each model described in Part Two.

Documenting Learning with Assessments

Assessment techniques will provide the teacher with a description of students' current learning, but how can the teacher know *how much* learning has occurred over the course of the unit? Summative assessments give a good picture of what students know and can do at the end of a unit but furnish no information about what students knew or could do when the unit started. Teachers can do two things to document the amount of student learning that occurs in a unit. They can give pre-assessments that cover the major learning outcomes in the unit and then compare the pre-assessments with summative assessments. Learning can then be defined as the difference in those scores, commonly called *gain*. Second, teachers can use continuous assessments to monitor student progress over the course of the unit, placing

EXHIBIT 7.6		Assessment strategies and timing with instructional models.			
MODEL	MAJOR LEARNING OUTCOMES	TRADITIONAL ASSESSMENTS	ALTERNATIVE ASSESSMENTS	TIMING	
Direct Instruction	1. Motor performance 2. Knowledge of rules 3. Fitness	Skills tests Written tests Fitness tests	Performance checklists	Formative Summative	
Personalized System for Instruction	1. Motor performance 2. Knowledge of rules 3. Fitness	Skills "quizzes" Written tests Fitness tests	Performance checklists Logs	Continuous	
Cooperative Learning	1. Knowledge of concepts 2. Team participation 3. Social development	Skills "quizzes" Written tests	Group projects Multimedia projects Portfolios Journals	Continuous Formative Summative	
Sport Education	1. Game performance and strategy 2. Team participation 3. "Duty role" performance	Written tests	Role-playing checklists Team projects Game performance summaries GPAI Journals	Formative Summative	
Peer Teaching	1. Motor performance 2. Cognitive knowledge (to teach) 3. Social development	Skills tests Written tests	Performance checklists Journals	Continuous Formative	
Inquiry Teaching	1. Higher-order knowledge, combined with 2. Movement skills	Written tests	Oral examinations Interviews Journals	Continuous Formative Summative	
Tactical Games	1. Tactical knowledge, combined with 2. Motor performance	Written tests	GPAI Performance checklists Oral examinations	Continuous Formative Summative	
Teaching Responsibility	1. Social development 2. Motor performance	Skills tests	Performance checklists Journals	Continuous Formative Summative	

students' performance scores in the many assessments on a chart to show the amount and the rate of learning.

Teachers can use individual or grouped information to document how much learning has occurred. Since each student will have her own set of scores over time, the teacher can examine those scores to determine if progress is being made. Teachers can group assessment data by calculating mean scores for subgroups of students (e.g., boys and girls, less skilled and more skilled) to note if any group shows a different amount of learning. If one group is found to be learning less, the teacher can change instructional strategies to help that group catch up. Assessment data can also be grouped for all students in the class by calculating a mean class score. Although class mean scores give the teacher an overall description of learning, they can often hide differences among individual students in the class.

Using Assessment for Planning and Teaching

Formative assessments can be used to monitor student learning as a unit progresses and to help teachers make informed decisions about changes that could be needed right away to increase the amount and rate of student learning. A performance assessment taken at the end of one day's lesson can provide a teacher with important information about whether students have learned the content well enough to move ahead in the sequence of learning activities, or whether students need more time and practice before moving on to new content, to more advanced content, or to playing competitive games. For example, if most students in a class score very low on a test of speedball rules and etiquette, it would not be a good decision to advance them to actual game play in the next lesson—even if the schedule calls for that to happen. This should clue teachers that further review of the rules and etiquette are needed and that they should conduct that review and a follow-up assessment in the next lesson, before students move on to game play.

Making Assessment Practical

Assessments give teachers valuable information about what students have learned in physical education. Teachers really need to know what students have learned, how much they have learned, and how well students have achieved stated standards and outcomes. As noted earlier, however, good assessments take time for teachers to prepare, for students to complete, and for teachers to review. And there is a catch—the time it takes a teacher to gather assessment information can take time away from student engagement with the content, actually reducing the amount of time available for learning.

Physical education teachers can use several strategies for gathering assessment information in a practical manner:

1. Do not attempt to assess all the potential learning outcomes in a unit. Identify the three or four most important outcomes, and assess only those.

2. Use assessments that have been successful for other teachers. Good examples of assessments can be found in many places, including the suggested readings listed at the end of this chapter.

3. Teach students to do self- and peer assessments when possible. Sometimes the teacher's main task is to watch students perform and count successful tries. Students from the middle grades and up can learn the key elements to look for and how to measure and quantify them; then they can watch and score their peers.

4. Use worksheets and scoring sheets to reduce paperwork. Many scores can be entered on one page rather than one score per page. This will help the teacher to be more organized and also allow for faster scoring of the results later.

5. Use small, continuous assessments as much as possible. Simply asking students to raise their hands if they have performed a learning task to a certain level can supply quick and valuable information.

6. Build assessment into learning tasks whenever possible. Rather than just asking students to practice for five minutes, direct them to count the number of "good" tries they make during those five minutes. Similarly, scoring sheets can be placed at learning stations, with students directed to fill in their scores before they move to the next station.

7. Use technology to gather, store, and analyze assessment information. Many physical education teachers are now using desktop computers, laptops, tablets, and even smartphones to enter assessments and store assessment data. This technology allows for convenient data entry and storage and facilitates generation and distribution of assessment reports to students, parents, and school administrators.

SUMMARY

Good physical education programs today must have standards for student learning, whether they are the NASPE standards or other stated goals and outcomes. Those standards represent the essential kinds of knowledge, skills, behaviors, and dispositions that students will achieve by participating in regular, quality physical education instruction. But having standards is not enough; good programs must be able to document the degree to which students have achieved those standards, based on relevant performance criteria—not just a teacher's opinion that learning has taken place. The key operation in documenting student achievement involves gathering assessment data and other information that can be used to establish a program's success or the need for improvement in certain areas. Because contemporary physical education programs strive for many and different outcomes and also because teachers want to help students to achieve in and out of school, assessment must take many forms and happen in many settings. The most effective teachers and programs are those that link standards, assessment, student learning, and continuous program improvement.

LEARNING ACTIVITIES

1. Go to the website of your state's department of education, and follow the links to its standards or other learning outcomes for physical education. (Your professor may need to assist you with this search.) For each standard/outcome, do two things. First, make a copy of the exact statement of that standard/outcome as it appears on the website. Second, use your own words to express the standard/outcome, as if you were trying to explain it to your students. That is, what would you tell students you are trying to teach that relates to this standard/outcome?
2. Select one standard/outcome in your state's list, and identify a grade level (preschool/lower elementary, upper elementary, middle school, high school). Next, identify a content unit you might teach for that grade level in order to address that standard/outcome (e.g., dance, fitness, team sports, basic throwing and catching). Explain how your content unit can lead students toward achieving that standard/outcome. What is the alignment between that unit and the standard/outcome you wish students to achieve?
3. Following up on the previous learning activity, state and explain three ways that students could demonstrate they have achieved the standard/outcome. One of those ways must occur *outside* the physical education program (e.g., in school, at home, in youth sports). Design an assessment strategy for each of those three ways of demonstrating the standard/outcome. Your strategy must include a description of the assessment task to be given to students, how it will be implemented, and how it will be scored. Finally, explain how you will use that assessment information to determine the degree to which your students have met the standard/outcome.
4. Select another standard/outcome, and repeat Activities 2 and 3.

TOPICS FOR REFLECTION & SMALL-GROUP DISCUSSION

1. Schools in the United States today are being held more accountable for student learning. Why do you think this is happening? What factors are contributing to this?
2. What is the relationship between learning standards (e.g., NASPE standards or your own state's standards) and assessment?
3. What kind of assessments did you have in physical education as a P-12 student? What was assessed, and how was that information used?
4. What are some of the reasons why meaningful assessments are difficult for physical education teachers to complete?
5. Pick any physical education content area that you know well (e.g., dance, team sports, fitness). What do you think is the best way to show that you have *really* learned that content?

SUGGESTED READINGS

- Lambert, L. T. (2007). *Standards-based assessment of student learning* (2nd ed.) Reston, VA: National Association for Sport and Physical Education.
- Lund, J. L., & Kirk, M. K. (2010). *Performance-based assessment for middle and high school physical education* (2nd ed.). Champaign, IL: Human Kinetics.
- Mitchell, S. A., & Oslin, J. L. (1999). *Assessment in games teaching*. Reston, VA: National Association for Sport and Physical Education.
- National Association for Sport and Physical Education. (2004). *Moving into the future: National standards for physical education* (2nd ed.). Reston, VA: Author.
- O'Sullivan, M. O., & Henniger, M. (2000). *Assessing student responsibility and teamwork*. Reston, VA: National Association for Sport and Physical Education.
- Schiemer, S. (2000). *Assessment strategies for elementary physical education*. Champaign, IL: Human Kinetics.