

Moving Into the Future

National Physical Education Standards: A Guide to Content and Assessment



**National
Association
for Sport &
Physical Education**

Developed by the
National Association for Sport and Physical Education



Boston, Massachusetts Burr Ridge, Illinois Dubuque, Iowa
Madison, Wisconsin New York, New York San Francisco, California St. Louis, Missouri



Copyright © 1995 by
The National Association for Sport and Physical Education
1900 Association Drive, Reston, Virginia 22091
All rights reserved.

NASPE is an association of the
American Alliance for Health, Physical Education, Recreation and Dance (AAHPERD)

WCB/McGraw-Hill

A Division of The McGraw-Hill Companies

Designed by Steve Reardon

International Standard Book Number: **0-8151-7338-5**



Appendix A: Assessment

Section 1: Overview

Two cornerstones of the current educational reform movement are accountability and assessment. Accountability refers to the responsibility of teachers for effective teaching and hence, to a certain degree, for student learning. Assessment refers to the process of testing and evaluating students to determine progress towards program goals.

Traditionally, the assessment model advocated for physical education involves

- formally stating the objectives of instruction (typically knowledge, psychomotor, and psycho-social objectives)
- preassessing students
- measuring achievement of objectives using valid and reliable tests during and after delivery of appropriate instructional activities
- evaluating student progress towards meeting the stated objectives in a formative framework using a criterion-referenced grading system.

Three primary principles underlay the traditional model of assessment (Wood, in press):

- **Establish appropriate instructional objectives.** Objectives for physical education can be found at the national, state, and local levels. The *NASPE Content Standards and Assessment Guide for School Physical Education* represents the most recent national standards for physical education. These standards reflect what a physically educated student should know and be able to do at each grade level. Rather than defining curriculum, these standards provide signposts for teachers and program accountability; although educators will find the standards useful in designing appropriate physical education curricula.
- **Use of appropriate tests to measure characteristics related to instructional objectives.** Assessment can take many forms. Informal assessment is used to enhance day-to-day instruction. It is usually nongraded, nonrecorded, and can range from simple observation of student performance and verbal checking for student understanding to nongraded pop quizzes. In contrast, formal assessment is used for gathering evidence to be used in formulating student grades, program evaluation, developing and revising program objectives, and for providing feedback for students. Formal assessment involves the administration of valid (tests that measure what they are supposed to measure) and reliable (measurement that is relatively error-free) assessment tools to evaluate the extent to which students have met the objectives of the program (see Section 2 below). The degree to which tests should be valid and reliable is in direct proportion to the importance of the decisions made based on test scores. For example, if sport skills tests are to account for 75 percent of the grade in physical education, the test should be chosen or developed with careful attention paid to validity and reliability. Traditionally, formal assessment has involved the use of sport skills tests, motor performance checklists and rating scales, knowledge tests, and subjective measures of psychosocial development. More recently; however, a number of alternative assessment strategies such as authentic assessment, and performance assessment are gaining

in popularity (see Section 3 below). The NASPE *Content Standards and Assessment Guide for School Physical Education* includes many assessment options and examples with each content standard. In addition, Section 4 below provides more detailed descriptions of assessment options used in the document. However, it must be emphasized that these options are only a sample of a broader range of available assessment tools. Detailed descriptions of how to develop valid and reliable checklists and rating scales, sport skills tests, and knowledge tests can be found in texts such as Safrit and Wood (1995) and Baumgartner and Jackson (1995). Other useful resources (see Appendix B) are Strand and Wilson (1993); Collins and Hodges (1978); Hensley, Morrow, and East (1990); McGee and Farrow (1987); and Schick (1981).

- **Development of an evaluation scheme that reflects attainment of the instructional objectives.** Grades in physical education should reflect attainment of stated instructional objectives. A teacher's grading plan should be a reflection of both the teacher's choice of objectives and the importance given to each objective. If a program emphasizes knowledge of physical fitness principles, performance in selected sports or dance and multicultural awareness, then the grading scheme should reflect each of those elements.

Section 4: Measurement Concepts

Two desirable characteristics of tests are validity and reliability. Validity refers to the appropriateness of test score interpretation, while reliability refers to test scores that are replicable or relatively free from error. The degree to which a test is valid and reliable reflects the degree of confidence that the test scores provide credible evidence regarding student performance.

The methodology for determining evidence for validity and reliability is diverse and ranges from logical analyses to complex statistical computations. For the public school physical education setting, however, a number of relatively easy methods provide adequate evidence. In addition, the degree to which tests should be valid and reliable is in direct proportion to the importance of the decisions made based on test scores.

Content-Related Evidence for Validity. For most tests the first steps in test development are (a) explicitly defining the student attributes to be measured (determined in relation to the instructional objectives) and (b) designing a test to measure the important attributes. If it can be shown that the attributes measured by the test reflect those attributes defined in step (a) then evidence has been presented for content validity. For traditional written tests (e.g., multiple choice tests) a table of specifications or blueprint that outlines the content to be covered by the test and the level of cognition (e.g., memorization, understanding and application of knowledge) required by students can be developed. Test questions are then written to reflect the content and behaviors outlined in the table of specifications. Content-related evidence for validity is assessed by matching the test questions to the table of specifications. If the test questions reflect the table of specifications then content validity is achieved. A similar process known as logical validity (Safrit and Wood, 1995) can be employed for tests of motor performance (e.g., a test of the tennis serve):

- The important elements of skill are listed.
- A test is constructed to measure the important elements of skill
- The test is scored such that higher scores reflect more proficient performance in the



defined elements

- If the elements examined by the test match the important elements outlined in the test purpose, a claim can be made for logical validity.

Reflection on the process of determining evidence for content-validity reveals that test construction is a systematic and goal directed process and not simply a matter of haphazardly developing test questions or elements. Achieving a valid test requires knowledge of the instructional objectives, contemplation of the important content and behaviors to be examined, careful construction of test questions and elements and examination of the degree to which the test questions and elements reflect the important content and behaviors.

Criterion-Related Evidence for Validity. While content-related evidence for validity is an appropriate starting point in the validation process, for some types of tests, most notably motor performance tests, additional evidence for validity can easily be determined. Criterion-related evidence for validity compares test scores with another test (the criterion) considered to be a better, although often less practical or feasible measure of the characteristic. For example, after determining the logical validity of a tennis serve test, a teacher administers the test to 30 students in his or her class. In addition, the school's tennis coach is asked to independently evaluate each student on their tennis serve. The coach's rating is the criterion or "gold standard" against which the field test is compared. Criterion-related evidence for validity is indicated if students who score well on the tennis serve test are also rated highly in tennis serving skill by the coach and if students who score low on the serve test are rated as less skilled in serving by the coach. A common method for quantifying this type of relationship is by computing the correlation coefficient between the tennis serve test scores and the coach's rating of each student. Correlation above 0.80 are acceptable for most purposes in a school setting (Safrit and Wood, 1995), although the acceptable magnitude depends in large part on the uses of the test scores.

Reliability. Reliability refers to the consistency of test scores over repeated testing or the relative freedom of scores from error. In physical education settings two types of evidence for reliability are important: *test-retest reliability* and *objectivity*. Test-retest reliability is the consistency of scores over time. That is, if you administer a test to students today and measure the same students using the same methods two days from now, the same scores should be obtained. Objectivity or rater reliability refers to the degree of consistency of scores when two different raters or the same rater on two occasions scores a group of students. Objectivity is a significant issue when scoring essay tests and written projects, observational instruments such as checklists and rating scales, and alternative assessments such as portfolios and event tasks. Clearly defined scoring criteria and significant practice using a scoring system lead to greater objectivity when using such assessments.

Section 3: Alternative Assessment

Traditionally, assessment specialists have advocated the use of standardized tests for assessing physical fitness, sport skills, knowledge, and psychosocial characteristics. Such tests have a high degree of validity and reliability and are accompanied by tables of norms or criterion-referenced standards. From the practitioner's point of view, however, such tests tend to be impractical and often fail to measure the instructional objectives of interest to teachers. Riding the current wave of educational reform with its emphasis on outcomes based education, integrated learning, and critical thinking skills, a number of alternative assessment strategies have captured the attention of teachers and administrators, eager to develop accountability systems to serve both the needs of students in a changing world and the demands of parents for responsibility in the classroom.

Alternative assessment can take many forms such as portfolios, discussions and debates, event tasks, case studies, student logs, and role playing. Such assessments are characterized by the following:

- tasks that directly examine the behavior the teacher wishes to measure
- a focus on product and quality of performance
- criterion-referenced scoring
- assessment of higher levels of learning
- student participation in development of the assessment and ownership of the final product
- assessment criteria that are given to students in advance (Bartz, Anderson-Robinson, & Hillman, 1994).

Rubrics are the scoring criteria by which student performance is judged. They are used most often with alternative assessments such as portfolios, event tasks, and student performance but can actually be used for other types of assessment as well. They are written by the teacher before instruction begins and shared with students as the unit or project is explained. Because students have the criteria very early, they have a standard by which they can judge their own performance, thus providing feedback during instruction.

When writing a rubric (see the following sample rubric), all important criteria should be addressed. If teachers are evaluating a project with several components, they may choose to write a single scoring rubric that addresses all the components (holistic rubric) or several rubrics to address each topic or goal (analytic rubric). Although the former is easier to score, the reliability is generally better with the latter. Either format is acceptable.

Start by writing down the components for the top level. There are different philosophies concerning the top level. One method is to write the ideal or optimal criteria so that very few, maybe only 1% or 2%, of your students will ever reach this. The reasoning behind this is that students will not have a ceiling effect and can always be striving to make their best better. Alternatively, write the top criteria at a level that better or "A" students would be able to achieve.

Most alternative assessments use rubrics with four to six levels. Each level is somewhat easier to achieve than the previous level as the standards are lower. The teacher must also decide what is the least acceptable level of performance. The levels of the rubric should be graduated between the upper and lower level.

Scoring Rubrics

Example of a scoring rubric:

Score 4 points if the student:

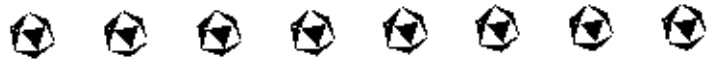
- Follows all directions and finishes all parts of the question
- Clearly answers the question so that others can understand
- Demonstrates that he or she really understands the information that is asked about and is not just giving related facts. In other words, the information is applied in some manner.
- Answers the question concisely, giving the best way to solve the problem presented
- Can apply the knowledge by showing connections between ideas and the real world, by comparing different ideas, and showing how the ideas work together

Score 3 points if the student:

- Follows the directions and finishes most of the parts of the question
- Clearly answers the question so that others can understand
- Demonstrates an understanding of the "big picture" related to the question but there may be a few little mistakes or wrong ideas

Score 2 points if the student:

- Follows some of the directions and finishes some parts of the question



- Answers the question clearly so that others can understand, but the answer may not be complete
- Demonstrates only a partial understanding of the knowledge and concepts necessary to complete the question

Score 1 point if the student:

- Understands only a small part of the information asked for in the question
- Answers only a small part of the question

Score 0 points if the student:

- Answers the question completely wrong or has nothing to do with the question

Blank

- Student did not give any answer at all

The level at which the rubric is written will depend upon several things; e.g., the length of time spent on the unit; age and ability level of student; amount of equipment available. Also, some of these assessments can be used for multiple purposes, depending on the teacher's goals. A teacher should start with a list of goals and concepts that students should know and then give lesser points for answers that are not complete or do not meet teacher expectations.

Alternative assessment is labeled "authentic" when the tasks are conducted in real-to-life contexts (Meyer, 1992). For example, to assess student learning in a physical fitness unit, traditional assessment involving written tests of knowledge and understanding or projects such as essays can be delivered. In contrast, an alternative assessment might consist of a group project to (a) assess the fitness needs of teachers and staff at the school, (b) design a custom physical fitness program for these individuals, (c) provide instruction on how to safely and effectively participate in fitness activities, and (d) monitor the individuals as they progress through the program. The teacher assumes the role of a facilitator to help students formulate appropriate questions and as a guide to finding appropriate answers. At each stage of the process, students present written and/or oral synopses and reflections of their challenges and successes, culminating in a portfolio that reflects both the products and process of student learning. Moreover, unlike more traditional types of assessment, feedback is given to students during the process so that learning takes place during the assessment. This type of assessment can be described as "alternative" to more traditional forms of assessment and as "authentic" because it assesses performance in a real-to-life rather than a contrived context.

The NASPE *Content Standards and Assessment Guide for School Physical Education* focuses on alternative assessment options because these types of assessment devices are not as common or as well articulated in physical education settings. This should not be interpreted as a condemnation of more traditional assessment devices. A balanced assessment strategy should employ assessment devices that best assess the instructional objectives of the physical education program.

From a measurement perspective, constructing and administering valid and reliable alternative assessments is a challenging task. Issues that require careful consideration include the following (Pierson and Beck, 1993):

- Are the tasks worthy of being assessed?
- How well do the tasks provided by the assessment match the proficiencies to be assessed (i.e., do the tasks measure what they are supposed to measure)?
- Is the end product clearly defined?
- Do the scoring rubrics, guidelines, or procedures accurately reflect the agreed-upon proficiencies?
- Are the criteria for administering and scoring the test precise and clear?
- Are the scoring rubrics precise enough to achieve a degree of inter-rater reliability?

- How will the assessment results be used (i.e., grading, improving instructional practices)?

Because many alternative assessments require elements of critical thinking, problem solving, and writing competency, developers of such tests should be aware of issues such as linguistic appropriateness and fairness or "opportunity to learn" so that some students are not disadvantaged by learning disabilities, lack of facilities or equipment, or language barriers (Baker, 1994). In addition, alternative assessment often requires teachers to critically analyze student work, a task for which teachers often are not trained and do not have time. Moreover, performance assessments can involve long-term group and individual projects that may require that teachers assume the role of facilitator rather than of instructor.

Alternative assessments provide an exciting assessment option for physical educators. Worthen (1993) provides 10 conditions important to a school's readiness to implement alternative assessment strategies:

- Desire for better assessment information.
- Indications that current assessment is creating negative side effects.
- Staff openness to innovations.
- Conceptual clarity about alternative assessment and its advantages and disadvantages.
- Assessment literacy.
- Clarity about desired student outcomes.
- Content or curricula ill-suited to traditional tests.
- School examples of alternative assessment.
- Willingness to critique assessment methods.
- Patron's and policy maker's openness to new forms of assessments.

The NASPE *Content Standards and Assessment Guide for School Physical Education* can assist schools with defining appropriate outcomes and providing examples of alternative assessments. NASPE invites you and your school to use this document as a springboard for maximizing the effectiveness of assessing physical education learning outcomes.

Section 4: Assessment Options

Some of the assessment options described in this document may not be familiar to readers. To assist those interested in using various assessment options, the following descriptions, examples, and hints for effective use are presented:

STUDENT PROJECT

Definition/Description:

Students engage in building a scenario, determining goals, planning a program of participation to achieve outcomes, and implementing the plan to the completion of the goals. Student projects provide for a range of strategies and results including the following: the application of the processes of data collection, goal setting, planning, analysis, decision making, problem-solving; development and application of skill and knowledge to real-life situations to solve problems or create "new" interventions to reach personal goals; may include multiple objectives or outcomes; combine multiple assessment options (logs, journals, reports); student autonomy in choosing procedures and reaching conclusions; solo or multiple students; multiple resources; changes in status, behaviors, conditions; authenticity; performance products; flexibility of time (complexity of task determines time); and, integration of multiple content areas, concepts, and applications.



Projects are assigned at the beginning of student learning sequences (units, courses) and are integrated with instruction. Criteria for assessment projects are presented at the initiation of the assignment.

Example: (Suggested Application: Grade Levels 10, 12 Standards #1, 2, 3, 4, 5, 6, 7)

Student analyzes interests, desires, capabilities and commitment for engaging in a physical activity of his or her choice. Using personal data about health and motor fitness status, create a plan for developing skills and fitness necessary for participation in a series of episodes in the chosen activity. Include a plan for skill improvement, practice, fitness conditioning, securing equipment, time, facilities, instruction, finances, and other people with whom to participate. The teacher may be used as a resource for planning and locating appropriate opportunities for participation. The project may be used to assess achievement of several other outcomes. The student will provide the following evidence of completion: a videotape of one episode in the series; a log of participation throughout the series; a written summary of the experience reflecting feelings about success, benefits, enjoyment, and potential for lifelong participation.

Criteria for Assessment:

1. Analyzes personal fitness status to plan skill and fitness development.
2. Applies basic skills and movement concepts to perform proficiently and creatively.
3. Applies principles of training to improve skill and fitness.
4. Uses resources to solve problems that enhance or limit participation.
5. Reflects on the benefits, enjoyment, and challenges that result from participation in physical activity.

Scoring:

- Exemplary: Successfully completes the series of episodes, demonstrating synthesis of skill, knowledge and attitudes to plan and perform proficiently and creatively, and assumes all the responsibilities of reporting the results.
- Acceptable: Completes the series of episodes, demonstrating application of skills, concepts and attitudes to perform basic skills with competence required for the pleasurable performance and assumes all the responsibilities of reporting the results.
- Needs Improvement: Planning, participation, and reporting reflect insufficient skill and knowledge.
- Unacceptable: Planning, participation, and reporting are incomplete.

Report Form:

- Checklist of multiple forms
- Written participation plan
- Records of fitness assessment (health and motor)
- Log of participation
- Videotape of performance
- Summary report

Hints for Development and Use:

- Experience with a variety of teaching styles will increase the successful direction and completion of student projects.
- Small projects that give students increasing amounts of responsibility for their own learning should be given in the early grades to prepare them for the complexity of this assessment.
- Criteria for assessment and scoring procedures are explained to the student at the beginning of the project.
- Multiple scorers that include community experts, cross-disciplinary, and multiple grade levels might be used.

- Pilot field testing of this option should be completed before results of this assessment are used for purposes of promotion or graduation.
- Student projects permit a high degree of individuality.
- Scoring rubrics will be necessary for each component of the project.
- Before assigning projects, evidence should be available which indicates mastery of basic concepts and skills necessary for successful completion.
- The element of choice provided by projects enables acceptance of learning of essential skills, concepts, and practices in a way that is important to the student.

STUDENT LOG

Definition/Description:

Students record performance of specific behaviors over a period of time that identifies products, time intervals, decisions/choices, and reflections. Recorded items should indicate critical factors relative to expected results. Information may show performance changes, sequence of behaviors, choices, feelings, documentation of conditions, progress, process, and/or regularity of participation. Logs may be kept by individual students, small groups, or whole classes. Information can be used to justify program change, make predictions, and in combination with other assessment options.

Example: (Suggested application: Grade Levels 2, 4, 6 Standard #3)

Students are asked to share their involvement in physical activities during nonschool times. Entries are made on a collective class wall chart to indicate what, when, and where activities are performed and the number of students who participated. Anecdotal records are made by the teacher for individual students indicating what influences the student's participation including the level of success/enjoyment expressed, student likes or dislikes, involvement with family and friends.

WALL CHART:

DATE	ACTIVITY	# OF PARTICIPANTS	WHERE	TIME	COMMENTS:
4/1	Rope jump	6	Yard	30 min	Double Dutch I don't like when the rope is too fast

Criteria for Assessment:

1. Activities selected have potential for vigorous physical activity.
2. Students use available opportunities to be involved in physical activity at least three times a week.

Scoring:

- Exemplary: Participates more than three times a week
- Acceptable: Participates at least three times a week
- Needs Improvement: Participates less than three times a week
- Unacceptable: No participation in vigorous activity

Report Form: Oral report to teacher



Hints for Development and Use:

- Critical factors about information to be collected determine the data to be recorded on the log.
- Reporting forms must be simple for quick recording.
- Data collected provides information that may be used for other assessment options or instructional strategies (e.g., learning curve).
- Keeping a log is a motivational tool.
- A log provides a tangible record of progress for both the student and the teacher.
- Individual logs promote student responsibility for personal learning.

STUDENT JOURNAL

Definition/Description:

Student record of participation, results, responses to, feelings, perceptions, or reflections about actual happenings or results. Entries, made at regular intervals over time, may serve as indicators of success, failure, benefits, or other intangible products of participation. Entries are not viewed as right or wrong since they are reflections about personal performance including social and psychological perspectives. Students may describe both positive and negative behavior. Journal entries are used to summarize, compare and contrast like and unlike experiences, provide opportunity for self-analysis of personal meaning and quality of participation, record behavior adjustments, compare results of other assessment options including conditions which contribute, enhance, or limit participation, and as a resource of suggestions for change. Journal entries can be reviewed to determine how a student processes both internal and external information about his or her performance.

Example: (Suggested Application: Grade Level 10 Standard # 7)

During an adventure education experience (i.e., ropes course, climbing a wall, nature hikes, camping, canoeing), record in a journal the feelings and thoughts experienced throughout.

Criteria for Assessment:

1. Analyzes and expresses feelings about physical activity.
2. Identifies evidence of success, challenge, and enjoyment present in the activity.
3. Explains challenge that adventure activities provide.
4. Describes the positive effects friends and companions bring to this experience.

Scoring:

- Exemplary: Expresses feelings of personal participation and in sharing it with friends.
- Acceptable: Identifies feelings of personal participation
- Needs improvement: Has difficulty expressing feelings about participation.
- Unacceptable: Does not make journal entries.

Reporting Form: Composition notebook

Hints for Development and Use:

- Writing should be used in simple episodes at first.
- Writings should not be judged as right or wrong.
- Student writings should be protected as private information to be shared only by the student or with student permission.

- Students might be given a choice to determine who may read the journal.
- Students may be asked to use the criteria for assessment to summarize the journal for scoring and reporting purposes.
- Summary reports might be used as an interdisciplinary assessment including communication arts and social studies.

PARENTAL REPORT

Definition/Description:

Record of student regularity, progress, process or product of participation that has been verified by the parent(s). The report may include verification by signature of a student's recorded report or by anecdotal comments of the parent or person who has observed the out-of-class performance.

Example: (Suggested application: Grade Level K, 2 Standard #3)

An anecdotal record is kept of observations over a period of time about the physical activities of children while under the supervision of parents or guardians and during nonschool time. The report should include play choices, purposeful practice, formal activities (sports clubs, dance lessons), or family activities. The report may be in the form of a log or a journal.

Criteria for Assessment:

1. Participates in a variety of physical activities.
2. Participates in health-enhancing activities regularly.

Scoring:

- Exemplary: Participates in a variety of activities more than three times per week.
- Acceptable: Participates in health-enhancing activities at least three times a week.
- Needs improvement: Participates in health enhancing physical activities less than three times a week.
- Unacceptable: Does not participate in physical activities outside of the school program.

Reporting Form:

(NAME OF SCHOOL)

PHYSICAL EDUCATION - PARENTAL REPORT FORM

Student Name _____

Parent Name _____

Describe the involvement of the student named above in physical activities each day during the week of _____. Include formal/informal, family/individual activities, the length of time spent, and any comments that indicate the level of performance.



Hints for Development and Use:

- Directions for parents should be written and reviewed by a small group of parents before involving parents of all students.
- Results may indicate special individual needs, weaknesses, or interests.
- Results may indicate parents with special interests that may be a useful resource.
- Results may indicate a need for instructional emphasis to enhance special community interests that are attractive to students.
- Parental involvement may help to stimulate more interest in planning physical activities in which the whole family can participate.

INTERVIEW

Definition/Description:

One-to-one discussion with a planned sequence of questions to obtain information (e.g., cognitive, affective, statistical). Most often thought of as teacher-to-student interviews for the purposes of obtaining information on student thoughts, feelings, and understandings. However, student-to-student or student-to-persons in the community interviews may be used for such purposes as analyzing activity patterns or computing frequency of exercise.

Example:

Interview two persons from each of the following age groups: 20 to 30 years, 40 to 50 years, 65+ years. Plan your interview questions in order to determine the physical activity for each individual. Use the information obtained from the interview to evaluate physical activity patterns to determine if each person is taking advantage of the physiological, psychological, and social benefits of physical activity.

Hints for Development and Use:

- Questions for the interview should be planned and sequential.
- The interview should be planned with a clear purpose in mind.
- Teacher-to-student interviews offer excellent opportunities for encouragement as well as for feedback information.
- When asking questions of students, be patient in waiting for the student to articulate his or her response.
- Remember, the purpose of the interview is to obtain information, not to impose your opinions.
- Student-to-others interviews are usually for obtaining information. Teacher-to-student interviews provide insight into the student's personal feelings, perceptions of strengths as well as student comprehension.

PEER OBSERVATION

Definition/Description:

The observation of students by other students to assess competence in performance of skill and demonstration of selected critical elements of skill. It is most often used for the observation of critical elements that lead to a mature execution of a particular skill. Informal peer observation is used throughout teaching to help students evaluate progress toward the goal (e.g., inclusion of all components for a gymnastics routine or creative dance, correct pathway of travel in response

to task). Peer observation feedback includes verbal discussion, verbal response, thumbs up or thumbs down, and written feedback. Videotape is a helpful support technology for peer observation.

Example:

Students observe for critical elements in the preparatory phase of a designated skill. For example, student "A" throws a ball toward a target five times using the overhand throw. Student "B" (the observer) focuses on the critical element stated by the teacher as the focus of the exercise (e.g., opposite foot forward, side to the target). The peer observer signals thumbs up if the critical element is correctly executed.

Hints for Development and Use:

- Criteria for assessment must be clear to the observer (i.e., clearly stated and understood). Teachers could provide students with explicit diagrams or written descriptions of the correct tasks.
- Students will need practice in observing. Being able to focus on a component rather than watching the total action is a learned skill. Students should not be expected to observe more than one criteria at a time.
- The feedback system should be simple for the observer (e.g., plus or minus, thumbs up or thumbs down).
- Peer observation is a valuable tool if criteria for assessment are understood by both the observer and the performer.
- **Reminder:** The purpose of peer observation is to assist with student learning, not to degrade students or their performance.

SELF-ASSESSMENT

Definition/Description:

The student assessing personal progress as opposed to being assessed by the teacher or by other students. Self-assessments include rating scales for levels of performance, participation, recording performance scores (e.g., distance, accuracy), summary reports after a series of assessment tasks (e.g., dribbling, throwing for accuracy and distance, jump shooting; physical fitness profiles), and questionnaires of likes and dislikes in activities. Self-assessment is a part of logs, journals, and portfolios as students evaluate personal performance or progress toward goals.

Example:

Students are involved in a ropes course activity unit in which they are asked to work together to accomplish group goals. At the completion of the day's activity, each student rates personal progress toward the following: assuming a leadership role, assuming a follower role, assisting someone in the group, working cooperatively with the total group.

Hints for Development and Use:

- Trust your students; they will be very honest in their assessments of themselves.
- Self-assessments provide teachers with insight into students' perceptions and self-concepts as they write summary statements and comments regarding personal performance.
- Students may need guidance in assessing themselves based on past performance and personal goals as opposed to comparisons with standards or others.
- Self-assessment can be a valuable tool in helping students accept responsibility for personal



activity and fitness, as well as in setting goals for the development of healthy lifestyle patterns.

WRITTEN TEST

Definition/Description:

Written tests encompass multiple choice, true/false, matching, essay, short answer, and fill-in-the-blank test formats traditionally used to examine knowledge, comprehension, application, analysis, synthesis, and evaluation of the knowledge base in physical education. Broadly speaking, such tests could include other test formats such as oral examinations and examinations that use drawings or pictures to elicit student responses. Written tests are commonly used for short quizzes or for longer formal examinations.

Example:

Physical fitness is a recurring theme throughout public school physical education. The following questions are provided as examples of the various types of questions that might be included on written tests at various grade levels:

K - Primary Grades: Teacher develops a pictorial display of people engaged in various activities ranging from jogging to watching television. Student are examined orally by asking such questions as:

- Teacher points to two pictures, one with a physically active person (jogger) and one with an inactive person (someone reading) and asks the student to choose the picture that shows the person with the fastest beating heart.
- Teacher points to a weight lifter and an jogger and asks which activity is best for building strong muscles.

Middle School to High School: Multiple choice, true/false, matching, and essay questions can be employed when students are able to read at an appropriate level.

- Sue decided to develop her own personal exercise program. Knowing that you were enrolled in PE 101 Physical Fitness Training she asked your advice in developing her program. Referring to the five steps that should be considered in developing personal exercise programs that we discussed in class, what advice would you give to Sue?
- Distinguish between health-related and athletic-related physical fitness. In addition, provide examples of the types of fitness tests used to measure each.
- Which of the equations listed below is commonly used to estimate maximum heart rate?
 - a. $220 - \text{weight}$
 - b. $220 - \text{age}$
 - c. $220 - \text{resting heart rate}$
 - d. $220 + \text{resting heart rate}$
- T/F The optimal training program for developing muscular strength is to lift light weights over many repetitions.
- To improve the physical fitness component known as _____ you should practice stretching exercises.

Hints for Development and Use:

- Written tests should represent a comprehensive sample of the content and behaviors outlined in the instructional objectives. Before writing test questions, a table of specifications or test blueprint outlining the content and behaviors to be elicited by the test can be developed to ensure



- that the test represents appropriate content and behaviors.
- Construction of meaningful test questions that reflect appropriate content and elicit appropriate levels of cognition requires practice. Rules for developing test questions can be found in measurement texts such as Saifit and Wood (1995) and Baumgartner and Jackson (1995).
- Students should be made aware of the general content areas covered by a test and the types of questions (e.g., multiple choice or essay) that will be presented.
- Written tests should be developed for the reading and comprehension level of students. Oral presentation of test questions for students in the lower grades or for students with reading difficulties can facilitate the test-taking process.
- Detailed scoring keys are required for essay and short answer tests to increase fairness and objectivity in scoring.

GROUP PROJECT

Definition/Description:

An assessment project completed by several students working cooperatively. As opposed to an event task that can be completed in a single class period, the group project usually takes more than one class period to complete and may include time spent outside of class. Group projects may be performance-based (e.g., presentation of dance, creation of a new game) or involve class presentation of results, displays, wall charts.

Example:

Students are placed in groups of five to six members and are asked to role-play the following scenario: The City Little League coach has asked your group to serve as assistant coaches this season. Specifically, the coach asks you to do a presentation of throwing for varying distances. Each group is instructed to prepare a presentation to include: oral presentation of skills, demonstration of skills, and audiovisuals. Each group member must be part of the presentation. Presentations will be made to the class.

Hints for Development and Use:

- Keep the groups small (four may be the maximum number for some classes).
- Provide guidelines for items that must be included in the project, especially performance skills.
- Within a class, some groups will require assistance to move beyond discussion of detail; others will require assistance to attend to detail.
- The project or game must be complete in itself (i.e., in **final** form).
- The scoring rubric should be shared with students in advance so they are aware of the assessment criteria.
- Keep the project as simple as possible.

PORTFOLIO

Definition/Description:

Portfolios are collections of a student's work assembled over time (Feuer and Fulton, 1993). They include various pieces of evidence documenting student achievement of a goal. Portfolios have been used by artists and models for many years to demonstrate their best work. The focus in student portfolios is on:

- Student thinking



- Growth over time
- Views of oneself as a learner
- Problem-solving

Portfolios should be rated according to their ability to convey to the teacher that the student has met the goals for the class. This should be done concisely. The cover letter that accompanies the portfolio clearly explains why these particular pieces were included (instead of other types of documentation) and how the teacher's goals were met by the student. Criteria should take into consideration the age of the student, the time given for portfolio development, and whether or not the portfolio is being used to fulfill requirements for other teachers and classes.

Examples of portfolio components: (*Note: Not all of these would be included in one portfolio*)

Goal: To learn to play soccer:

- Evidence of playing on a recreational team (e.g., certificate)
- Journal of student successes as skill improved
- Videotape of game play
- A list of the drills practiced three times a week
- A letter from a soccer coach
- A brief review of key rules
- A critique of a soccer game watched discussing various offensive and defensive strategies observed.
- A practice log (e.g., footwork skills) to show improvement in the number that can be done in 30 second time frames
- Chart on skills

Goal: Fitness Development:

- Research paper on training components
- Training log to document distances
- Student diary that describes how he or she felt after each training episode
- A chart that records resting heart rate over time
- Summary paper of the experience
- Log of strength and flexibility activities
- "Before and after" body fat assessments
- An outline of a training program created before the fitness unit began
- A participant's number from a local road race competition
- Pictures
- Journal excerpts
- Graphs of improvement
- The original fitness plan and a corrected plan allowing adjustments for improvement

Examples of portfolio content in sport skill classes:

- Student written self-evaluation of current skill level and individual goals
- Daily log of class activities and individual performance, suggestions for future work
- On-going self, and peer analysis of skill performance and playing performance (checklist, rating scale, criteria-referenced tasks, videotape)
- Based on self-analysis, students select or design appropriate practice program and complete schedule
- Graphs that chart daily practice/play performance on major skills
- Documentation of practice, informal game play and/or organized competition outside of class time
- Set up, conduct, and participate in a tournament or meet, keep group and personal sta-

tistics

- Write a newspaper article reporting on the class tournament as if you were a sports reporter
- Complete and record a play-by-play commentary as if you were a radio or television sports announcer
- Design optional routines according to class rules and practice for competition
- Interview a successful competitor or coach about his or her development and write an article
- Read an autobiography about a successful competitor and write a review
- Write a final analysis of your skill and playing or performing ability, assign your own grade and give rationale
- Write an essay entitled "What I learned and accomplished in _____ . . . and what I learned about myself in the process"
- Generate a reference list of instructional materials
- Attend a clinic, workshop or presentation and write a review

Hints for Development and Use:

- The first step in creating a portfolio involves a teacher decision — determining the desired student outcome or goal. Once this is established and clearly stated, the student can begin gathering evidence that will demonstrate that the goal has been met.
- Portfolios are gathered over time. They may cover 6 to 8 weeks of instruction. Therefore, students need to have adequate time to assemble their information as well as adequate resources. The portfolio is often considered to be assessment in progress. The student has the option of working on components, changing them in ways that students feel make the evidence stronger so that they have achieved the desired goal.
- If a student included every bit of evidence, the portfolio could become massive. Grading these unlimited portfolios would be a monumental if not impossible task. Unlimited portfolios are referred to as working portfolios. For the assessment portfolio, a teacher may limit the portfolio to a certain number of pieces (e.g., seven) and then have the student write a cover letter explaining why the various items were included. Although a teacher may suggest portfolio components or areas to be addressed, the ultimate decision of what is to be included should be left to the student.
- One of the key components of alternative assessment is that the student has ownership and full choice about what is to be included. For instance, let's say a student was learning a psychomotor skill or sport. The teacher thought that a practice log, listing time spent working on the skill was necessary to demonstrate the practice and learning. Instead the student turns in a video showing early game play and game play following practice. The latter piece would also demonstrate skill practice, even though a log of practice time had not been kept. Another example might be if a swimmer included a certificate from the Red Cross documenting completion of a life saving class.
- The teacher must be prepared to accept some unusual forms of documentation in the portfolio. This is not to say that the student always makes the best choice of selection of what to include. A good rubric should explain clearly the intent of the teacher goal, without limiting student creativity to meet the criteria.
- Portfolios are best used when learning involves a complex idea or skill. When many facets and components of learning are involved, it is easier for students to create a portfolio.
- Contents of a portfolio could include the following:
 - Report of a group project
 - Excerpts from a student's daily journal
 - Photo, model, or sketch made by the student
 - Notes from an interview or conference
 - Teacher-completed checklists
 - Copies of awards or prizes
 - Video, audio, and computer-generated examples of student work





ROLE PLAYING

Definition/Description:

Students are given a scenario and then asked to simulate the characters they portray or act out the situation that has been set for them. These dialogues can be written or verbalized. Students have the opportunity to portray real world situations. Students are required to use reasoning and problem solving to deal with the reality of the experience as it unfolds.

Examples:

1. Following a wall climbing class, let a student assume the role of Sir Edmond Hillary and discuss how he/she felt as the summit of Mt. Everest grew closer. Have a student "reporter" interview Sir Edmond.

Note: Rubrics have not been written for these role play examples (see Appendix A, Section 3 for a description of a scoring rubric). The rubric would address the points the teacher felt were important and tried to address with students. The question could have been written differently, and the teacher could have looked at what a student had learned about wall climbing techniques. Another version of the same scenario could have Sir Edmond discussing his training program and conditioning before beginning the trip. Still another could have looked at safety issues. A teacher could have combined two or more of these scenarios into a much broader look at learning.

2. You are at a meeting of the Major League Baseball owners. The topic on the table is league mascots. Some owners want the Indians and Braves to change their symbols. Create the dialogue for the meeting.

3. You are at a local basketball game. Your best friend's sister is on the opposing team. Whenever the team shoots free throws, another friend whistles and makes disruptive noises. Your friend also has been booing officials and is making derogatory remarks to the other team. What would you say to this friend? Create a dialogue for the conversation.

4. You are at a baseball game (Note: you could substitute tennis or racquetball) with a friend from France. This person knows a lot about soccer and basketball, but not much about baseball. How would you explain the rules so that your friend could enjoy the game?

Hints for Development and Use:

- Role playing lends itself easily to cognitive or affective domain assessment. The teacher must be careful to set the stage so that students know what is expected.
- Performing role plays in front of peers may be embarrassing or very difficult for some students. Introduce this activity to students gradually, beginning in small groups without an audience and progressing *gradually* to "performances" for the class. Begin with issues that are familiar to or popular with students before progressing to more personal scenarios.
- The teacher should accept the responses of students in a nonjudgmental way. Students should reflect, paraphrase, and summarize responses that have been given verbally.

EVENT TASK

Definition/Description:

An event task is a performance task that can be completed within 50 minutes. The task is writ-

ten broadly enough (it is loosely structured) so that there are multiple solutions or many possible correct answers. It should be engaging (capture the interest of the students) and replicate or simulate a real world experience.

Example:

Your group has been asked to organize a half-time show for the local basketball team using the various skills you have developed this year in physical education. Identify the skills you are going to present and create a routine to be performed. Make a list of equipment that you will need, if any. Choose your skills wisely so that everyone in the group is able to perform his or her part of the routine. You may include both individual and group stunts and skills. The performance will last 5 to 6 minutes.

Scoring Criteria:

Score the task a 4 if the students:

- Show variety in their choice of skills
- Choose skills of an appropriate level of difficulty for each member of the group
- Have sufficient skills to last the required time
- Demonstrate an understanding of equipment needs (e.g., tumbling mats, items to juggle, dance music)
- All members of the group are active throughout the routine

Score the task a 3 if the students:

- Show variety in their choice of skills
- Most of the skills are of an appropriate level of difficulty for each member, however, some members are not proficient in everything
- There are enough skills to last the required time
- Equipment needs for the routine are complete and appropriate
- Most members of the group are active throughout the routine, however one or two stand without activity for brief periods

Score the task a 2 if the students:

- Show variety in their choice of skills
- Most of the skills are of an appropriate level of difficulty for each member, however, there are several errors caused by students trying to do things that are too difficult
- The routine lasts about 4 1/2 minutes
- Equipment needs for the routine are complete and appropriate
- Most of the time everyone is active, however, there are places in the routine that are dominated by the most skilled performers

Score the task a 1 if the students:

- Use the same skills repetitively
- Have frequent errors caused by doing things that are too difficult
- The routine is very short
- Some of the necessary equipment has been omitted from the list
- Are active much of the time but the whole group does little together
- Do not have a good flow to the routine

Score the task a 0 if the students:

- Do things individually instead of having designed a routine

Score the task blank if the students:

- Do not do a routine



Hints for Development and Use:

A performance event should:

- Solicit responses representing all proficiency levels if possible
- Require minimal to no interaction with the person administering the test
- Result in individual responses that can be scored
- Allow variable student grouping (two to four students)
- Be completed within 50 minutes
- Have all materials necessary for successful completion readily available to students

OBSERVATION - TEACHER/STUDENT

Definition/Description:

Teacher observation is the most utilized form of assessment in physical education. Teachers observe students on a regular basis as part of the instructional process. Teacher observation can also be used more systematically to provide objective data on student performance to collect information on the instructional process or to evaluate students. All students or a sample of students representing different skill levels can be assessed. There are many tools that teachers can use to record observational data including anecdotal records, checklists, rating scales, or scoring rubrics. All of these tools can be used whether in live observation or with video analysis. They can be used by the teacher or by the student in peer assessment or self-assessment. The value of the information recorded in each of these cases is enhanced if teachers have a clear idea of what they are looking for in their observations and attend to issues related to the reliability and validity of the data they collect.

Anecdotal Record: The teacher establishes broad categories of concern and then takes notes on everything that is observed in relation to those categories. Notes are usually kept in the form of a log or running description and are usually nonevaluative.

Example: The teacher is interested in observing how students integrate mainstreamed students into their group work. During a group assignment the teacher records the interaction of mainstreamed students in various groups.

Observational record: Xavia did not go to the area of the gym with the rest of her group. One student from the group noticed that she was not part of the group and went over to get her. Thomas was included immediately in the group and given the assignment by the group to work with one of the more highly skilled students. Jose just wandered the gym until the teacher brought him to his assigned group.

Checklist: Checklists are usually used to determine whether a student meets or does not meet a particular criteria. The teacher checks off when a student meets the criteria.

Example: The teacher checks off when a student meets each of the criteria for an overhand throw.

Observational record:

- Steps with the opposite foot
- Sequential rotation
- Elbow out and back
- Forearm lag

Rating Scale: Rating scales are used to determine the degree to which identified criteria have been met in an observation. The teacher designates a) the important components of behavior to



be assessed and b) the levels of performance desired to be discriminated.

Example: The teacher has decided to systematically observe the degree to which students comply with class rules. Several students each class period are targeted for observation.

Observational record: For each class rule the teacher records one of the following levels of performance:

- 1= No incidence of noncompliance are observed
- 2= One incidence of noncompliance is observed
- 3= More than one incidence of noncompliance are observed but student is primarily compliant
- 4= Student is more often noncompliant than compliant

Scoring Rubric: Scoring rubrics are in one sense rating scales that use multiple criteria simultaneously. The teacher establishes desired levels of performance and then defines the criteria that are essential for each of those levels of performance. See Appendix A, Section 3 for additional information regarding scoring rubrics.

Example: The teacher wants to evaluate offensive game play in a three vs. three soccer game using a goalie.

Observational record:

Level 1:

- Passes ahead of receiver
- Moves into a logical empty space when receiving
- Maintains possession until defended

Level 2:

- Usually passes ahead of the receiver
- Moves into empty space to receive but not always logically in relation to the defense
- Does not always wait for defense before getting rid of the ball

Level 3:

- Passes are more often not ahead of the receiver
- Does not move into an empty space to receive the ball
- The decision to pass is not made in relation to the defender

Video analysis: All of the techniques for observation included above can be used either live or with the help of video analysis. Teachers and students can use video in time outside of class or during class. Video analysis has the advantage of slow motion capability that is useful for many skills that are too fast to be accurately analyzed via live observation. Video analysis also has the advantage of being replayed to increase the reliability of the observation. Teachers who do not have the luxury of evaluating each student during class time can use video analysis to get very complete and accurate information of the performance of individuals. Video analysis is also extremely useful to assess the degree to which program objectives and goals have been accomplished. Often teachers who see their classes on videotape are able to see a lot more when the responsibility for instruction is removed.

Example 1: The teacher videotapes all of the final games of a sport unit and establishes a scoring rubric based on the degree to which students use the individual skills of the game, know and use the rules, and use the game strategies that have been taught.

Example 2: The teacher sets up a corner of the gym for videotaping of a skill that has been taught. Each student takes a turn videotaping his or her performance in this skill. The teacher goes



through the tape to indicate the counter number associated with the location on the tape of each student's performance. The teacher puts the tape in the school media center and requires each student to use a rating scale to assess his or her own performance and to determine what aspects of performance need improvement.

Hints for Development and Use:

- Establish specific criteria for observation ahead of time so that you know what you are looking for and how you are going to observe.
- Choose a method of recording and plan an observation record that is easy to use.
- Do not try to look at too many criteria at one time.
- A lot of useful information can be obtained by sampling students rather than trying to observe every student every period. If all students need to be observed, choose several a day as your focus. If you are using observation for instructional feedback or program assessment, choose students who are likely to be at different levels of performance as what you are observing.
- When you need a high degree of reliability and validity for your assessment, practice using your criteria until you can use them consistently in a variety of contexts.