



PE•METRICS™

**Assessing
National Standards 1 - 6
In Secondary School**

PE•METRICS™

Assessing National Standards 1-6 In Secondary School

PE Metrics: Assessing National Standards 1-6 In Secondary School

Published by

体育韵律学：美国国家中学标准 1-6 评价



National Association for
Sport and Physical Education

an association of the American Alliance for Health,
Physical Education, Recreation and Dance

NASPE Sets the Standard

1900 Association Drive, Reston, Va. 20191 (703) 476-3410 www.naspeinfo.org

© 2011, National Association for Sport and Physical Education, an association of the American Alliance for Health, Physical Education, Recreation and Dance. All rights reserved.

PE Metrics™ is a trademark of the National Association for Sport and Physical Education, an association of the American Alliance for Health, Physical Education, Recreation and Dance.

Reproducing this work in any form by any electronic, mechanical or other means — including photocopying or storing in any information-retrieval system — is expressly forbidden without written permission from the publisher. To request permission to reprint or copy portions of this work, visit www.naspeinfo.org or e-mail naspepermissions@aahperd.org.

To order more copies of this book (stock # 304-10504):

Web: www.naspeinfo.org

E-mail: customerservice@aahperd.org

Phone: (800) 321-0789; (412) 741-1288 outside the United States

Fax: (412) 741-0609

Mail: AAHPERD Publications Fulfillment Center, P.O. Box 1020, Sewickley, PA 15143-1020

ISBN: 978-0-88314-960-7

Printed in the United States

Acknowledgements

Assessment Task Force

Marybell Avery
Ben Dyson
Jennifer L. Fiset
Connie Fox
Marian Franck
Kim C. Graber
Judith H. Placek, chair
Judith Rink
Lori Williams
Weimo Zhu

Technical Assistance

Marco Boscolo
Connie Fox
Pamela MacFarlane
Sara Khosravi Nasr
Youngsik Park
Eng Wah Teo
Weimo Zhu

Cognitive Assessment Writing Team

Suzan Ayers	Connie Fox
Bonnie Mohnsen	Marian Franck
Sarajane Quinn	Lois Mauch
Lori Williams	Jim DeLine
	Jennifer Reeves

NASPE Staff

Cheryl Richardson, Senior Director of Programs
De Raynes, Program Manager for Physical Education (ret.)
Joe McGavin, Publication Director

NASPE extends its appreciation to the many professionals who served as project administrators and coders in the gathering of thousands of pieces of individual student data for the secondary-level Standard 1 assessments. That list includes: Ann Arns, Marybell Avery, Deb Berkey, Tim Bott, Janet Brandt, Mary Buddemeier, Kenny Cope, Hugh Ferry, Jen Fiset, Connie Fox, Ted France, Marian Franck, Ann Marie Gallo, Jennifer Gorecki, Kim Graber, Tina Hall, Mark Harris, Leslie Hicks, Korey Hill, Christine Hopper, Diedre Jago, Michelle Leising, Stephanie Little, Sue Long, Susan Lynn, Koji Matsushima, Marguerite McDonald, Shauna McGhie, Doug McLeren, Marcia Motz, Diane Mozen, Jamie O'Conner, Karen Pagnano-Richardson, Ben Peasley, Penny Portman, De Raynes, Judy Rink, Gabe Romero, Anne Catherine Sullivan, Phillip Thomas, Julia Valley, Lori Williams, Rachel Williams and Russ Williams.

In addition, NASPE extends its appreciation to more than 100 middle school and high school physical educators, from all across the nation in every type of school and situation, who found time in their limited class hours to administer one or more of the Standards 2-6 cognitive tests. Their participation yielded thousands of student tests, which provided a rich source of data for establishing validity, reliability, item discrimination and other advanced data analysis applications.

T able of Contents

Acknowledgements	ii
Introduction	1
Standards-Based Education.....	1
Standards-Based Accountability	2
Standards-Based Education & Accountability for Physical Education	2
Chapter 1: Overview of NASPE's PE Metrics Assessment Project	5
Performance Indicators & Assessment Development	5
Standard 1	5
Standards 2-6.....	6
Institutional Review Board Approval.....	6
Data Collectors	6
Standard 1	6
Standards 2-6	7
Process for Testing & Data Collection.....	7
Pre-Pilot Data Collection.....	7
Standard 1	7
Standards 2-6	7
Pilot Data Collection.....	8
Standard 1	8
Standards 2-6	8
National Data Collection	8
Standard 1	8
Standards 2-6	9
Summary	9
Chapter 2: Using PE Metrics Assessments.....	11
Using the Assessments in This Book	11
1. Teachers.....	11
Using PE Metrics for Formative & Summative Assessment	11
Sampling Students to Collect Assessment Data.....	12
Grading.....	12
Reports to Parents.....	13
Teacher Preparation for Using PE Metrics Materials	13
2. District Administrators.....	13
3. Local, State & National Policymakers	14
4. Researchers	14
Misusing Assessment	14

Chapter 3: Administering & Scoring PE Metrics Assessments	17
Standard 1	17
Procedures for Administering the Assessments & Collecting Data.....	17
Administering the Assessments.....	17
Standard 1	17
Preparation.....	17
Video-Recording	18
Safety Considerations.....	18
Warm-Up & Practice.....	18
Test Administration.....	18
Coding Protocols.....	19
Preview/view the video record before trying to score.....	19
Practice Scoring	19
Scoring.....	19
Scoring Sheet Information	20
Standards 2–6	20
Chapter 4: Psychometric Quality of Assessments	21
New Concepts, Theory & Techniques Employed	21
Online Bank of Assessments	21
Item Response Theory.....	21
Test Equating.....	22
Calibration of Standard 1 Assessments	23
Samples & Data Collection.....	24
Summary of Results.....	24
Calibration of Standards 2–6 Assessments.....	24
Summary of Results.....	25
Using PE Metrics Appropriately.....	25
Conclusion	25
Table 1: Grade 8 Standard 1 Descriptive Statistics & Frequencies	27
Table 2: High School Standard 1 Descriptive Statistics & Frequencies	29
Table 3: Summary of IRT Calibration of Middle Assessments	30
Table 4: Summary of IRT Calibration of High School Assessments	31
Table 5: Sample of Size Summary by Grade, Standards & Form	31
Table 6: Item Analysis Criteria	32
Table 7: Item Statistics of Common Items and Form A for Grade 8, Standard 2.....	33
Table 8: Item Statistics of Common Items and Form A for Grade 8, Standards 3 & 4.....	35
Table 9: Item Statistics of Common Items and Form A for Grade 8, Standards 5 & 6.....	37
Table 10: Item Statistics of Common Items and Form A for High School, Standard 2	39
Table 11: Item Statistics of Common Items and Form A for High School, Standards 3 & 4	41
Table 12: Item Statistics of Common Items and Form A for High School, Standards 5 & 6	43
Chapter 5: Frequently Asked Questions About PE Metrics.....	45
It's so overwhelming to think about all this assessment. How can I get started?.....	45
I don't teach many of the activities in the PE Metrics Standard 1 assessments. How can they help me?	45
In regard to the Standards 2-6 tests, what if I don't teach everything in the test bank?.....	46
What if I want to assess my students at levels other than the end of grade 8 or high school?	46
I have collected a lot of assessment scores from my students. What can I do with the data?	46
How can I use the results of the PE Metrics assessments to assign a grade?.....	46

I don't have the exact equipment — such as a team handball — or space that the assessments call for. Can I substitute another ball or reduce the size of the space?.....	46
PE Metrics recommends video-recording the assessments for Standard 1. Why should I video-record the assessments?	47
I don't have time to video my classes and then view the recordings. Can I just score students "live" as they complete the assessments in class?	47
What are the procedures used in Standard 1 outlined in the chapter titled Administering & Scoring PE Metrics Assessments, and why are they important?.....	47
How can I supervise my class, operate the video camera and ensure meaningful participation in my large classes and still conduct Standard 1 assessments?.....	47
I have my students only twice a week for one year. How can I use PE Metrics assessments without taking up half of my teaching time?.....	48
How am I to use the PE Metrics Web site?	48
What is the PE Metrics National Database and why should I consider contributing to it?	48
Standard 1 Assessments	49
Grade 8 Assessments	51
Archery	51
Badminton	55
Floor Hockey	60
Folk Dance	64
Line Dance	68
Pickleball	71
Soccer	75
Softball	79
Team Handball	83
Traverse Climbing.....	87
Ultimate Frisbee®	91
Volleyball.....	95
High School Assessments	101
Basketball.....	101
Bowling	106
Canoeing.....	111
Flag Football	116
Golf	120
Soccer	126
Swimming	130
Tennis	134
Volleyball.....	139
Wall Climbing	143
Weight Training	148
Standards 2–6 Performance Descriptors & Sample Questions.....	153
Grade 8	155
Standard 2	155
Standards 3 & 4	160
Standards 5 & 6	165

High School

Standard 2	169
Standards 3 & 4	174
Standards 5 & 6	180

Appendix A: National Standards for Physical Education and Secondary-Level Indicators & Assessment Tools.. 185

Resources 192

I ntroduction

Standards-Based Education

Since the late 1980s, education reform in the United States has been driven largely by setting academic standards that describe what students should know and be able to do, and developing accountability systems for measuring student achievement of those standards. Terms such as “standards-based education,” “standards-based education reform” and “standards movement” are common in today’s education vernacular.

Standards and assessment have been pivotal themes in recent reform efforts, cutting across much of the federal legislation that Congress has passed to improve education for all students. Standards-based education began in 1994, when Congress passed Goals 2000: The Educate America Act. That established a framework for identifying world-class academic standards, measuring student progress and providing support for any student needing help in meeting the standards. Goals 2000 codified eight national education goals:

1. School readiness.
2. School completion.
3. Student achievement & citizenship.
4. Teacher education & professional development.
5. Mathematics and science excellence.
6. Adult literacy & lifelong learning.
7. Alcohol- and drug-free schools.
8. Parental participation.

Since Goals 2000, Congress has twice amended and/or reauthorized the Elementary and Secondary School Education Act of 1965 (ESEA). The 1994 ESEA reauthorization, dubbed the Improving America’s Schools Act, focused on changing the way education is delivered, encouraging comprehensive school reform, upgrading instruction and professional development to align with high standards, strengthening accountability and promoting the coordination of resources to improve education for all children.

The most recent ESEA reauthorization, the No Child Left Behind Act of 2001 (NCLB), prescribes increased accountability for states, school districts and schools, greater choice for parents and students (particularly those attending low-performing schools), more flexibility for state and local education agencies in using federal education dollars and a stronger emphasis on reading, especially for the youngest children. Together, those federal laws have established many of the principles of standards-based reform, including the expectation that all students will attain high standards of academic excellence.

Standards-based education requires clear, measurable standards for all students. Standards and benchmarks identify what students should know and be able to do as they progress through school, and should be written so that they are developmentally appropriate and relevant to future education and employment needs. They also should be written so that all students are capable of achieving them, and so that talented students will exceed them.

Standards are meant to be anchors, aligning curriculum, instruction and assessment. That explains the emergence of the terms “standards-based curriculum,” “standards-based instruction” and “standards-based assessment,” as well as the holistic term “standards-based education.” By design, standards-based education lends itself to accountability.

Standards-Based Accountability

In years past, accountability in education was determined by measuring school inputs and processes such as funding levels, curriculum offerings and resources and regulation compliance. After national and state standards emerged, policymakers began to shift the focus of accountability to student outcomes. Now, policymakers emphasize student learning and achievement outcomes to gauge the success of state and local education efforts. This trend in education reform is known as standards-based accountability. Systems have been and continue to be put into place to hold states, districts, schools and teachers accountable for students’ performance.

Standards-based accountability systems use criterion-referenced performance standards rather than norm-referenced rankings. A standards-based system measures each student against a concrete standard, instead of measuring how well a student performs when compared to others.

Once standards are established, the next step in building a standards-based accountability system is aligning curriculum to the standards. The knowledge or skills that students must acquire to meet the standards must be defined, and criterion-referenced assessments must be established to determine the extent to which students meet the standards. The alignment of instruction and formative assessment of the standards must follow.

Standards-Based Education & Accountability for Physical Education

The National Association for Sport and Physical Education (NASPE) has established itself as a leader in standards-based education. In 1995, NASPE published *Moving Into the Future: National Standards for Physical Education*, followed by a second edition in 2004.

“Physical activity is critical to the development and maintenance of good health, state NASPE’s National Standards. “The goal of physical education is to develop physically educated individuals who have the knowledge, skills and confidence to enjoy a lifetime of healthful physical activity” (NASPE, 2004).

NASPE defines a physically educated person as someone who:

1. Demonstrates competency in motor skills and movement patterns needed to perform a variety of physical activities.
2. Demonstrates understanding of movement concepts, principles, strategies and tactics as they apply to the learning and performance of physical activities.
3. Participates regularly in physical activity.
4. Achieves and maintains a health-enhancing level of physical fitness.
5. Exhibits responsible personal and social behavior that respects self and others in physical activity settings.
6. Values physical activity for health, enjoyment, challenge, self-expression and/or social interaction.

The National Standards are presented in grade-level ranges representing K-2, 3-5, 6-8 and 9-12, so that the ranges are consistent with children’s and youths’ developmental patterns, that they reflect organizational patterns in public school settings and that they align with other content areas.

Each grade range contains two sections: student expectations and sample performance outcomes. Student expectations reflect what students should know and be able to do at the end of each grade-level range (e.g., grade 8). Sample performance outcomes are examples of student behavior at each grade-level range that demonstrate progress toward meeting the standards.

Until now, nationally tested assessments and rubrics to measure student achievement of the National Standards and benchmarks have been the missing elements of standards-based physical education. With the publication of PE Metrics for elementary, now followed by PE Metrics for middle and high school, NASPE has closed the gap.

PE Metrics provides valid and reliable standards-based assessments and rubrics (“NASPE assessments”) to measure student achievement of the National Standards by high school graduation and appropriate progress at four other grade-level ranges. The assessments in this book give evidence of learning through student work/performance, and their rubrics describe the quality of the work/performance. With PE Metrics, teachers and schools have the ability to report student progress on each standard. The advantage to that approach is that it provides students, teachers and parents with highly specific information.

It’s critical to align curriculum, instruction and assessments with one another and with state and national standards. To provide students with a truly standards-based physical education, teachers must be certain that the material on which students are being assessed aligns with what’s being taught in class.

Assessment as part of the instruction process constitutes much more than evaluation and accountability. Teachers should use a variety of techniques, including NASPE assessments, as part of regular classroom instruction. Assessment integrated with instruction (e.g., pre-assessment and formative assessment) is imperative for maximizing student learning and success. It’s equally important to use pre-assessment and formative assessment to communicate important skills and knowledge to students and to prepare students properly for summative assessment.

For accountability purposes, it isn’t necessary to assess all students on every task; one can use a random sample of student performance to guide curriculum development and/or to report on programmatic success to school, district and/or state leaders.

Whether one uses PE Metrics assessments as part of the instruction process (e.g., pre-assessment and formative assessment) or for evaluation and accountability, they are scientifically valid and useful tools for measuring student achievement.

Chapter 1

Overview of NASPE's PE Metrics Assessment Project

The National Association for Sport and Physical Education (NASPE) is committed to the tenets of high-quality physical education, which include the opportunity for all students to learn through supportive policies and environment (e.g., certified teachers, adequate facilities and equipment), National Standards, high-quality curriculum, appropriate instruction practices, and student and program assessment. NASPE continues to develop tools to help schools, districts and states implement assessments that measure student progress toward state and national physical education standards. As the only national association for physical education, NASPE feels a strong obligation to develop valid and reliable assessments that teachers throughout the nation can use.

In March 1999, NASPE assembled a “think tank” of university and public school professionals to consider how to best advance K-12 physical education. The group was charged with recommending a plan of action to confront the barriers to high-quality physical education. The priority recommendation was to develop performance indicators and practical assessments to evaluate student progress toward the National Standards for Physical Education.

In January 2000, NASPE created what became known as its Assessment Task Force, made up of experts in curriculum and instruction and measurement and evaluation, researchers, teacher educators, K-12 physical education teachers, administrators and students. NASPE charged the task force with developing performance indicators that corresponded to the National Standards at each grade-level range (K-2, 3-5, 6-8 and 9-12), as well as assessments for each indicator. The performance indicators were not meant to be a comprehensive set of all possible skills and knowledge that students should master in a specific grade-level range, but rather samples of performance outcomes that could be expected within that grade-level range.

Performance Indicators & Assessment Development

Just as it was not feasible to identify all possible performance indicators, it became clear to the task force that it would not be feasible to write all possible assessments for each indicator. Ultimately, the task force identified a broad range of performance indicators and wrote a variety of assessments to measure student skills and knowledge. The examples selected and presented in this book and CD-ROM provide templates of good assessments that can serve to guide teachers, districts and states in developing additional assessments.

Standard 1

Initially, task force members were paired with content experts, and writing the assessments occurred over a period of several years. The draft performance indicators and assessments that they developed were introduced to 220 professionals attending a session at the 2001 American Alliance for Health, Physical Education, Recreation and Dance (AAHPERD) National Convention. NASPE continued to revise the indicators and assessments and to solicit member feedback during the development stages at AAHPERD National Conventions each year, as well as at other state and regional conferences.

Standards 2-6

Once the Standard 1 assessments were written and in the process of being pilot-tested, the task force focused on developing assessments for Standards 2-6. After writing performance indicators, the task force originally developed a series of authentic assessments for Standards 2-6. As early drafts of the assessments were tested, however, it became clear that the task force would not be able to develop authentic assessments (e.g., determining student ability to cooperate in class) that were reliable and valid. Therefore, the task force decided to use written tests (multiple-choice) to measure learning outcomes for Standards 2-6. These tests were written by small committees of content experts, under the direction of task force members.

The test-writing committees developed tables of specifications to ensure that the questions were written at the appropriate level of difficulty and reading comprehension and were linked to the content within each performance indicator. Rather than developing five separate tests (one for each standard), the committees grouped the questions into three broad concept areas (Standard 2; Standards 3 & 4; and Standards 5 & 6). The committees wrote three forms (A, B & C), with 40 multiple-choice items each for the grade 8 and high school levels.

Institutional Review Board Approval

In preparing for the data collection, NASPE obtained Institutional Review Board (IRB) approval for using human subjects in a research study from the University of Illinois at Urbana-Champaign (UIUC) and Northern Illinois University (NIU). NASPE obtained IRB approval in 2003, and it has been renewed each year. Although the IRB document was written to enable project administrators from different universities to use the same consent documents, the project administrators were advised to make sure that UIUC's IRB approval would satisfy the IRB requirements at their home institutions. In the end, all project administrators submitted joint agreements with UIUC or were covered under NIU's IRB. NASPE obtained an additional IRB from NIU for administering Standards 2-6 tests.

Following data collection for the elementary project, NIU granted an IRB exemption for the Standards 2-6 multiple-choice tests for the national data collection, because of the nature of the anonymous multiple-choice test format and because these types of assessments are commonly conducted in physical education classes.

Data Collectors

Standard 1

NASPE trained teacher educators to supervise the administration of the pilot and national data collection for Standard 1. These project administrators were chosen based on four criteria:

1. They had established contacts with teachers and administrators in the public schools through their placements of pre-service students for field experience.
2. They had experience with video-recording student performances.
3. They had knowledge of the appropriate execution of motor skills.
4. Their research backgrounds meant that they understood the importance of adhering strictly to testing protocols, including obtaining school district permission and parent/student consent for testing, as well as following the exact assessment description and instructions, including equipment, site preparation, safety and video-recording.

Project administrators were recruited through personal contacts made by NASPE staff and task force members. They received six hours of training that included a project history and overview, roles and responsibilities (including gaining entry, informed consent, video-recording and coding), protocols for video-recording and conducting the assessments, typical video-recording problems and guidelines for coding the video records. Training took place at AAHPERD National Conventions (2004, 2005, 2006 & 2007) and at regional and district conferences (2004, 2007).

Standards 2-6

Task force members and NASPE staff recruited secondary physical education teachers from around the country to administer the tests to their students. The tests, detailed instructions, consent forms and Scantron answer sheets were sent to the teachers and subsequently returned to the researcher at NIU for initial analysis. In addition to administering the tests, teachers were asked to review questions for content and readability and to provide feedback to the task force.

Process for Testing & Data Collection

Standard 1 assessments and Standards 2-6 written tests were subjected to the following sequence of testing:

- Pre-pilot data collection.
- Pilot data collection.
- National data collection.

Because Standard 1 measures skill performance, testing involved video-recording each assessment, then coding the videos to produce scores for each aspect of students' performance. Because of the seasonal nature of instruction in physical education (e.g., restrictions imposed by inclement weather, some units taught only in the spring or fall), pilot and national data collection took longer than anticipated. Because Standards 2-6 assessments are paper-and-pencil tests that measure students' cognitive knowledge, they were easier to administer, required fewer personnel and were completed more quickly than the Standard 1 assessments.

Pre-Pilot Data Collection

Standard 1

All assessments were pre-piloted by at least one secondary teacher and 20 students. Teachers provided feedback on the instructions to teachers and students, camera placement and level of difficulty for the assessment task. The task force revised the assessments based on that feedback, as well as on feedback obtained during AAHPERD National Convention sessions. During the pre-pilot phase, NASPE reduced the number of National Standards from seven to six. As a result, in 2005, the task force modified both the performance indicators and assessments to meet the new standards and grade-level ranges. Accordingly, the task force collected additional data and pre-pilot feedback for some of the Standard 1 assessments.

Standards 2-6

Three forms of each test at each grade level were administered to a minimum of 20 students in December 2008, and teachers provided feedback on readability, clarity and ease of administration. The task force revised questions based on teacher feedback and on conventional data analysis (item

difficulty and discrimination). Based on the analysis, the task force identified the 10 strongest questions, which were used as common questions on each form of the three tests (A, B & C) to be used for pilot testing.

Pilot Data Collection

Standard 1

Project administrators subsequently tested the assessments on a larger scale to further ensure their appropriateness and to collect preliminary data on discrimination, validity and reliability. The task force collected data from various sections of the country through an extensive network of project administrators.

At least 40 students completed each pilot assessment. Project administrators collected parent/student/teacher consent forms, worked with a secondary physical education teacher to administer the assessments and video-recorded, coded and provided feedback on the assessments. The data were analyzed to ensure that the assessments were appropriately difficult and that they revealed meaningful differences among students. The task force revised the assessments as needed, based on data analysis and the feedback from teachers and project administrators.

Standards 2-6

The task force and NASPE staff recruited teachers from around the United States at AAHPERD and district conventions, and from personal contacts to conduct one or more cognitive assessments, and 200 students completed each of the three forms of the Standards 2-6 test in January and February 2009. Data analysis identified a problem with discrimination and difficulty levels with a number of questions, so the task force revised the problematic questions. All the questions were pilot-tested again in October 2009 with a smaller number of students. The resulting data were analyzed, and the best questions were used to create two forms (A & B) of the test for each grade level for the national data collection.

National Data Collection

Standard 1

Data collection for the Standard 1 assessments began in February 2005 and continued through spring 2010. (The IRB approval from UIUC continued to cover national data collection.) NASPE requested an additional IRB, which NIU approved each year from 2007 to 2010 to cover a subset of data collectors not named originally in the UIUC IRB. (See Table 1 on p. 27 for the list of assessments for which national data were collected.)

The project administrators' goal was to collect data from a minimum of 200 students for every assessment task. Two common tasks of medium difficulty were identified at each grade-level range, and students at each grade-level range completed at least one of the two common tasks. The common tasks for Standard 1 secondary assessments were:

- Grade 8: soccer 3 on 2 and volleyball.
- High School: volleyball and weight training.

The data from the common tasks were used in the research process to allow scores to be placed on a common scale. A common scale is important for being able to equate assessments across and within grade-level ranges and to create an assessment bank. These topics are explored further in the Psychometric Quality of Assessments chapter.

Standards 2-6

National data collection for the secondary grade levels occurred in spring 2010, with 400 students completing each form of the test (A & B) at each grade level (grade 8 and high school). Data analysis was conducted in June 2010.

Summary

NASPE's Assessment Task Force has followed a long and difficult path in creating the PE Metrics assessments. In all, the process took more than 10 years of concerted, ongoing effort by the entire task force and writing team members. Task force members believe that this careful and extensive process has resulted in valid and reliable assessments that are useful for teachers, administrators, teacher educators and researchers in both physical education and the wider field of education in general. The task force views its work as a first step in NASPE's ongoing effort to create and then add to a body of assessments for our field.

