

Physical and Health Education (9—12) Item Bank Specifications



Hillsborough County Public Schools

Florida Department of Education

Race to the Top

2012

This project was developed as part of the Florida Department of Education's Race to the Top Initiative.

As a statewide initiative, teachers from districts throughout the state contributed to the development of these materials.

The following districts worked in partnership with HCPS to contribute to the success of the project:

DeSoto
Duval
Escambia
Hendry
Lake
Leon
Pinellas
Polk
Osceola
St. Lucie

NOTE: The contents of these item bank specifications were developed under a grant from the U.S. Department of Education. However, those contents do not necessarily represent the policy of the U.S. Department of Education and you should not assume endorsement by the Federal Government.

Table of Contents

Introduction	4
Scope of the <i>Specifications</i>	6
Overall Considerations	6
Universal Design.....	7
Criteria for Physical and Health Education Assessment Items	8
Item Style and Format.....	8
General Item Guidelines	8
Item Types.....	8
Multiple Choice Items	8
Constructed Response Items	11
Performance Tasks.....	13
Rubrics.....	15
Media in Test Items.....	16
Types	16
Item Difficulty and Cognitive Complexity	18
Item Difficulty	18
Cognitive Complexity	18
Low Complexity.....	19
Moderate Complexity	19
High Complexity	20
Guide to the Item Specifications.....	22
What is an item specification?	22
Physical and Health Education Benchmark Classification System.....	22
Common Core Benchmark Classification System -Math	23
Item Specifications Template and Definitions	24
Guidelines for Item Writers	25
Review Procedures	26
Appendix A: Physical and Health Education Standards and Benchmarks	28
Appendix B: Item Writing Template	34
Appendix C: Item Writing and Review Checklist.....	35
Appendix D: Item Review Cover Sheet	37

Introduction

On August 24, 2010, the U.S. Department of Education announced Florida a winner of the federal Race to the Top Phase 2 competition. An important component of Florida's winning application focused on the "Standards and Assessments" assurance area, including the creation of high-quality, balanced assessments. At a national level, assessments generally focus on "traditional" content areas such as Reading, Math, Science and Social Studies. However, Florida has recognized the need for high quality, standards based assessments in other content areas, such as Physical Education and Health Education. Through this grant, seven projects were awarded to Local Education Agencies (LEAs) individually or in partnerships to create high-quality assessments for hard-to-measure content areas. The seven projects and the district awardees are as follows:

Project Letter	Hard To Measure Content Area	District Awardee
A	Physical and Health Education, K-8	Miami-Dade
B	Physical and Health Education, 9-12	Hillsborough
C	Performing Arts	Polk
D	Performing Arts	Polk
E	Visual Arts	Miami-Dade
F	World Languages	Duval
G	Career and Technical Education	Hillsborough

Within each of these Content Areas, the Florida Department of Education sought and included recommendations from educators across the state, including the Florida Organization of Instructional Leaders (FOIL) and a State Advisory Committee on District-Developed Student Assessments for Instructional Effectiveness (DDSAIE), made up of parents, teachers, and district-level administrators, to determine the appropriate scope of work for assessments for the hard-to-measure content areas. Based on these recommendations, teacher assignment and student enrollment data were analyzed to determine the courses that reach the greatest number of students and teachers in the hard-to-measure content areas. Hillsborough County partnered with Escambia, Leon, Duval, Polk, Osceola, Manatee, DeSoto, St. Lucie, and Hendry counties to develop the assessment items for Health and Physical Education.

This project is intended to provide Florida public and charter school districts with an extensive bank of assessment items that are high-quality, standards-based (NGSSS), field tested, and vetted by Florida educators. When the item bank and test platform are fully operational, Florida public and charter teachers and districts will have the ability to search the bank, export items, and generate customized assessments to meet their needs. In addition, there will be a public level of practice items available to students and parents that independent schools may access as well.

Hillsborough County is the fiscal agent for the development of Physical and Health Education item banks.

Test items and a subsequent item bank will be created for the following Physical and Health Education courses:

- Aerobics 1
- Aerobics 2
- Basketball
- Care and Prevention of Athletic Injuries
- Comprehensive Fitness
- Fitness Issues for Adolescence
- Fitness Lifestyle Design
- Health 1
- HOPE – PE Variation
- HOPE – Core
- Ind. and Dual Sports 1
- Ind. and Dual Sports 2
- Ind. and Dual Sports 3
- Outdoor Education
- Personal Fitness
- Power Weight Training 1
- Recreational Activities
- Team Sports 1
- Team Sports 2
- Volleyball 1
- Volleyball 2
- Weight Training 1
- Weight Training 2
- Weight Training 3

Origin and Purpose of this Document

The *Health and Physical Education Item Bank Specifications (Specifications)* document was created as a resource for item specification writers/reviewers and item writers/reviewers. The *Specifications* defines the content and format of the item specifications, test items, and overall test considerations. It addresses the content area of Health Education (Health) and Physical Education (PE) and the 24 courses included in this project, demonstrating the connection of test bank items with the NGSSS. The *Specifications* document serves to provide all stakeholders with information about the scope of the item bank for Health and PE.

Scope of the Specifications

The *Specifications* provides general and content area-specific guidelines for the development of all test items contained in the Physical and Health Education Item Bank.

The *Overall Considerations* section describes guidelines for test items in the Item bank, including Universal Design.

The *Criteria for the Health and Physical Education Assessment Items* section addresses the general guidelines used to develop multiple-choice items, constructed response items, and performance tasks. This section also covers the use of rubrics and media in items.

The *Item Difficulty and Cognitive Complexity* section addresses cognitive complexity levels as well as item difficulty and universal design.

The *Guide to the Item Specifications* section contains specific information about each standard and benchmark. It provides the template for the item specifications and defines the criteria for each item specification.

The *Guidelines for Item Writers* section includes general information for item writing.

The *Review Procedures for Health and Physical Education* provides an overview of the procedures and criteria that item reviewers utilize to review items.

Overall Considerations

This section of the specifications describes the guidelines that apply to all test items developed for the Physical and Health Education Item Bank. Overall considerations are broad item-development guidelines that should be addressed during the development of test items for the Physical and Health Education Item Bank.

1. Each multiple-choice item should be written to measure primarily one benchmark; however, other benchmarks may also be reflected in the item content.
2. Constructed response items and performance tasks may be written to reflect more than one benchmark.
3. All items, regardless of the item format, should be grade-level appropriate in terms of item difficulty, cognitive demands, and reading level.
4. The reading level of the items should aim to be two grade levels below the grade level of the students for whom they are intended, unless the student's reading ability is being assessed as part of the standard. If it is intended to be a reading item, it should be on grade level.

5. Items should not disadvantage or exhibit disrespect to anyone in regard to age, gender, race, ethnicity, language, religion, socioeconomic status, disability, occupation, or geographic region.
6. Some items may include an excerpt from stimulus material associated with several items in addition to the item stem.
7. Items should provide clear, concise, and complete instructions to students.
8. Each item should be written clearly and unambiguously to elicit the desired response.
9. Items may require the student to apply knowledge and skills described in NGSSS benchmarks from lower grades; however, the benchmarks from lower grades will not be assessed in isolation.
10. Items should provide all necessary components and information for students to listen, analyze, or use in order to respond to the items.
11. Items on the Health and Physical Education assessments should be written so that students are expected to select or provide the most accurate answer or appropriate response. Students should be allowed to listen to or view question stimuli for an appropriate number of times, or be allowed adequate time to prepare a performance when appropriate.
12. Constructed response and performance tasks must have a clear, concise rubric for grading.

Universal Design

The application of universal design principles helps develop assessments that are usable by the greatest number of test takers, including those with disabilities and non-native speakers of English. To support the goal of providing access to all students, the test items maximize readability, legibility, and compatibility with accommodations, and test item development includes a review for potential bias and sensitivity issues.

Universal design also implies that construct irrelevant features should be removed from items so that the item is truly assessing the intended construct and not some other feature of the item. It is expected that students in grades 9-12 taking the Physical and Health Education test items are able to read; however, for some benchmarks, their reading ability is not the construct being assessed. The students' scores on items with a complex reading passage will not represent their mastery of the course content but rather it will reflect their inability to read. Therefore, efforts will be made to minimize the amount and complexity of the text in test items, as applicable. If the intent of the benchmark is to assess the student's ability to read however, an appropriate text will be used.

In addition, because the Hard to Measure Content Areas item banks are assessing students' ability to apply the content and skills they have learned in class through activities that demonstrate the attainment of the required skills, the nature of the questions as well as the scored products will differ from those commonly used in standardized assessments. However, even with the inclusion of video components and audio components both as stimulus materials and as evidence for completed activities, the principles of universal design must be maintained so that students' scores reflect their knowledge of the constructs being assessed and not extraneous features of the task.

Internal and external reviewers revise items and tasks to allow for the widest possible range of student participation. Item writers must attend to the best practices suggested by universal design, including but not limited to

- reduction of wordiness
- avoidance of ambiguity
- selection of reader-friendly construction and terminology

- consistently applied concept names and graphic conventions

Universal design principles also inform decisions about test layout and design, including, but not limited to, type size, line length, spacing, and graphics.

Throughout the development process for the Florida Physical and Health Education Item Bank, these elements are carefully monitored. The review processes and field testing are used to ensure appropriateness, clarity, and fairness.

Criteria for Physical and Health Education Assessment Items

The Physical and Health Education Item Bank has three allowable item types: multiple choice, constructed response, and performance tasks.

Item Style and Format

This section presents stylistic guidelines and formatting directions for all items included in the Physical and Health Education Item bank. These guidelines should be followed while developing any test items for the Physical and Health Education Item Bank to ensure clarity and consistency. Specific guidelines for each item type are included in the appropriate “Item Type” section below.

General Item Guidelines

1. Items should be clear and concise, and they should use vocabulary and sentence structure appropriate for the assessed grade level.
2. The final sentence of any multiple choice item stem must be expressed as a question. If an item or task asks a question involving the word *not*, the word *not* should be emphasized by being bolded and written in all uppercase letters (e.g., “Which of the following is **NOT** an example of . . .”). As appropriate, other key words such as MOST, BEST, or LEAST should also be boldface and capitalized.
3. Masculine pronouns should not be used to refer to both sexes. Plural forms should be used whenever possible to avoid gender-specific pronouns (e.g., instead of “The student will make changes so that he . . .,” use “The students will make changes so that they . . .”).
4. An equal balance of male and female names should be used, including names representing different ethnic groups appropriate for Florida.
5. The comma should be used in a number greater than or equal to 1,000.
6. Acronyms are allowable only if they are commonly taught and used in the content area. (e.g., OSHA or FIFA)

Item Types

Multiple Choice Items

A **multiple choice** item is an objective test question which contains a question and four answer choices. One of the choices is the correct answer and the other choices are incorrect. The correct answer must be the only correct answer. Students who have mastered the content should be able to distinguish the correct answer from the other choices. However, the other choices must be plausible, as they are intended to draw students who have not yet mastered the concept being assessed.

Multiple choice items are designed to assess one benchmark. These items can be Low, Moderate, or High Complexity, based on Webb’s Depth of Knowledge Levels. These levels are further explained later in this document.

Terminology

In assessment, there are specific terms used to represent the different components of a multiple choice item.

Stimulus: what the student sees/hears/reads prior to the question. The stimulus gives the student necessary information to answer the question. The question is also included in the stimulus.

Stem: the question

Response Options: all of the provided answer choices

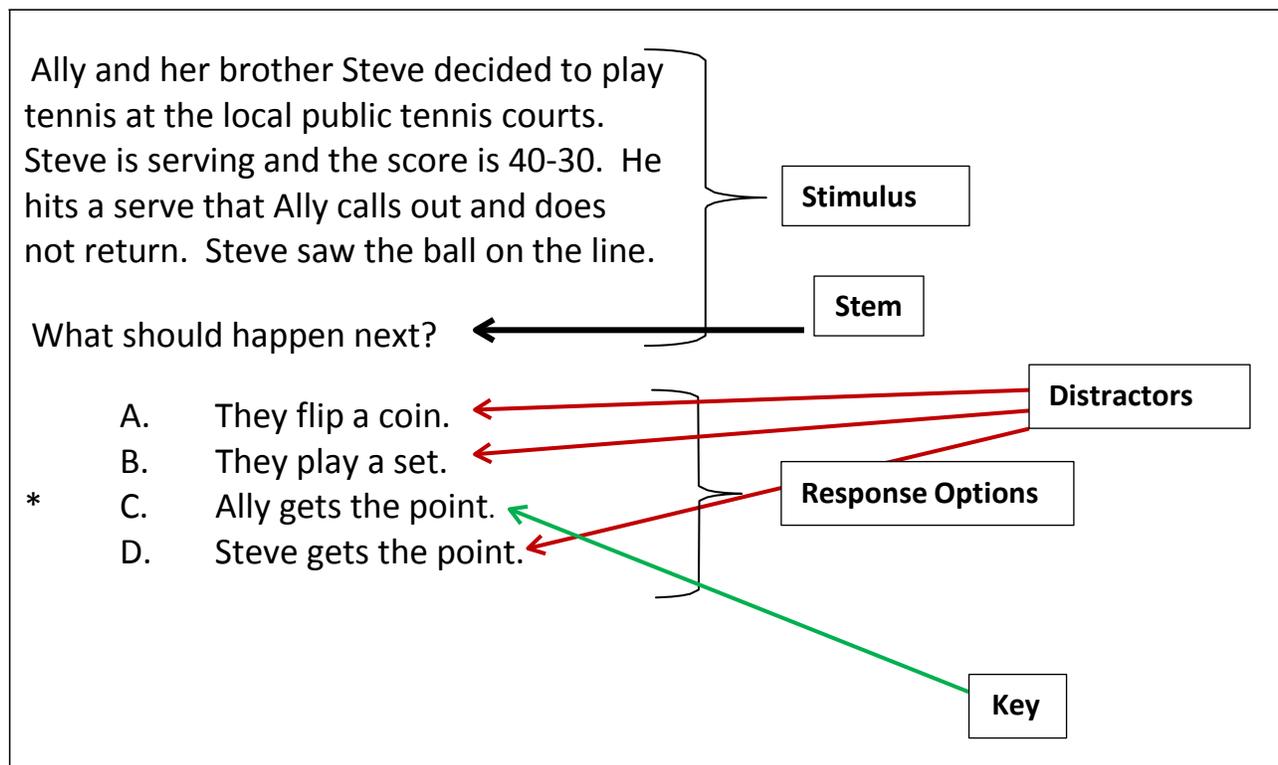
Key: the correct answer choice

Distractors: the incorrect but plausible answer choices

Sample Multiple Choice Item:

Course: Recreational Activities

Benchmark: PE.912.C.1.28 Interpret and apply the rules associated with specific course activities.



Criteria for Multiple Choice (MC) Items:

1. MC items should take approximately one minute of testing time per item to answer.
2. MC items are worth one point each.
3. MC items should have four response options
4. MC items should have only one correct answer.
5. MC items should have directed stems, not undirected stems, such as: Which of the following statements is true?" The student should be able to answer the question without looking at the response options first.
6. Response options should be parallel in reference to parts of speech (i.e., options may all be the same part of speech, may all be different parts of speech, or contain two pairs of the same part of speech).
7. Response options should be arranged by length; short to long or long to short, to allow for variety in the position of the answer key.
8. Answer keys should not be the only option that contains words or phrases used in the item stem.
9. Answer keys should not be the only positively or negatively worded options.
10. Distractors should relate to the context of the question. Distractors should be incorrect but plausible based on the topic of the question.
11. Outliers should be avoided because they are answer choices that clue or draw the student's attention away from the other answer choices. For example, one answer choice that is much longer than the other choices may be an outlier.
12. Response options should avoid using opposites. A test wise student can eliminate options this way.
13. Response options should not overlap. In other words, two options should not mean the same thing, but said in different words.
14. Response options should not include the following: *No change needed, Correct as is, None of the above, All of the above, Not enough information, Cannot be determined, etc.*
15. Response options should avoid absolutes such as "only" or "never."
16. Response options should avoid clang associations, or repeating sounds but not meaning. This makes it a trick question. For example, if the question stem has the word "bought" in it and the answer key has the word "thought" in it.
17. The answer key should not be the only response option to include words from the stem. All options or no options should include words from the stem.
18. If the response options are single words, they should not be capitalized unless they are proper nouns.
19. If the response options are phrases, they should not begin with a capital letter, nor should they have a period at the end.
20. If the response options are sentences, they should be capitalized and punctuated appropriately.
21. MC items should adhere to all the general guidelines listed above in addition to these guidelines.

Constructed Response Items

Constructed response items require students to provide a written response. These questions typically ask students to describe, discuss, explain, critique, or evaluate the scenario provided in the stimulus. Constructed response items have multiple correct responses. However, there may be specific content that is required in the response to receive full credit.

Constructed response items can assess one or more benchmarks and can range from low to high complexity. Constructed response items are scored using a rubric. Rubrics can include up to four score points.

Terminology

Listed below are the definitions of the specific terms used for the parts of a constructed response item along with a sample item with each of the terms identified.

Stimulus: the stimulus provides the students with the instructions for the task or it provides some type of input (such as a scenario, graphic, or video clip) to which the student must respond.

Rubric: a description of how to score the student's response

Constructed response items may require a short answer or a more extended response, such as an essay. Instructions are provided in the stimulus as to how much the student is required to produce.

The sample item below shows the format for a constructed response item with the item specific rubric. In this sample, the stimulus is presented first, followed by the rubric with the applicable score points. An extended constructed response item is worth from 1 to 4 points; however, a student who does not answer receives a score of "0".

Sample Constructed Response Item:

Course: Health 1

Benchmark: HE.912.C.2.6 Evaluate the impact of technology on personal, family, and community health.

Stimulus: In an effort to improve the health of Greentown residents, the mayor decided to purchase some new fitness technology to be used at the community center. He purchased some exergaming equipment and some new body composition scales. Residents who were interested in the program took a fitness assessment before the implementation of the new technology and then again after the technology had been in use for two months. The mayor found that 10% of the program participants moved into the healthy fitness zone in those two months. Evaluate the mayor's decision to purchase this equipment. Do you think he made the right decision? Why or why not? What suggestions do you have for the mayor on ways to use technology to further improve the overall health status of the residents of Greentown?

Rubric:

4 Points	Response shows a thorough evaluation of the impact of the technology on Greentown's residents' health. Response accurately evaluates the mayor's decision and includes the data mentioned in the stimulus in the response. Response supports his/her evaluation with many relevant examples and details. Response also includes realistic and appropriate suggestions for the mayor. Response is clear, but may include a few minor errors.
3 Points	Response shows an evaluation of the impact of the technology on Greentown's residents' health. Response evaluates the mayor's decision and may include the data mentioned in the stimulus. Response may include a few inaccuracies but is generally clear. Response supports his/her evaluation with some details and examples. Response includes some suggestions for the mayor, all of which may not be realistic or appropriate. Response may include some errors.
2 Points	Response shows a partial evaluation of the impact of the technology on Greentown's residents' health. Response evaluates the mayor's decision but with many inaccuracies. Response provides limited examples to support his/her evaluation. Response may not include the data mentioned in the stimulus. Response may include few suggestions, some of which are realist or appropriate. Response may contain many inaccuracies.
1 Point	Response shows a poor evaluation of the impact of the technology on Greentown's residents' health. Evaluation is mostly inaccurate and unclear. Response includes no relevant examples to support his/her evaluation and includes no realistic or appropriate suggestions for the mayor.

Criteria for Constructed Response (CR) Items

1. CR items should require a more complex response than a multiple-choice item permits.
2. CR items should not ask for a yes or no response.
3. CR items should contain enough information to focus the student on the task that must be accomplished.
4. CR items should indicate the amount and type of information that must be included in the response.
5. CR items should not ask for multiple repetitions of the same task to gain additional points.
6. CR items should not be used to assess lower level skills, unless the assessment task mirrors one required in the professional field.
7. CR items with two parts should not include score point dependencies, that is, achieving points on the second part of the item should not be dependent on giving a correct response to the first part of the item.
8. CR items should present a task that can be completed in the time allotted for the question type.
9. CR items should permit multiple solutions to the task presented in the prompt. Rubrics must provide for scoring of multiple solutions.
10. CR items should have rubrics that are clearly aligned to the task presented in the prompt.
11. CR items must be accompanied by a rubric or scoring guideline that lists the anticipated responses and the point value for each of the responses.
12. CR item rubric score point descriptions should clearly delineate what a student must do to earn each point.
13. CR item rubrics should not contain scoring rules that do not correspond to what is asked in the prompt.
14. Following field testing of the CR items, sample student papers should be included in the rubrics to provide student-generated exemplars for each of the score points.
15. CR items should adhere to all the general guidelines listed above in addition to these guidelines.

Performance Tasks

Performance tasks require students to produce a product or demonstrate required actions or behaviors. Performance tasks can range from low complexity to high complexity. Low complexity performance tasks may include the performance of rote behaviors required in particular Health and Physical Education courses. More complex performance tasks may include tasks that require the student to analyze the information presented in the stimulus and perform some action or actions as a response, or may require the student to create a product after following multiple steps.

Terminology

Listed below are the definitions of the specific terms used for the parts of a performance task along with a sample task with each of the terms identified.

Task: the directions that are stated prior to the student beginning the task

Rubric: the scoring guidelines for a student's response

Sample Performance Task:

Course: Volleyball 1

Benchmark: PE.912.M.1.34 Demonstrate use of the mechanical principles as they apply to specific course activities.

Task: Practice the overhand serve. Complete 20 overhand serves with proper form, appropriate application of force and correct contact with the ball. Criteria for a proper overhand serve is as follows:

- 1) Foot position: Student stands behind the end line, facing the net. The feet are staggered, pointing to the target.
- 2) Toss: The toss is consistent, and allows for same contact every time. The striking arm forms a bow and arrow and smoothly transitions to reach the head.
- 3) Contact: The contact occurs above and in front of the head with an open palm.
- 4) Follow Through: Follow through is complete, and the ball consistently goes over the net and towards the target.

Rubric:

4 Points	Student performs exemplary overhand serves 90-100% of the time, executing the 4 important components of a successful serve: foot position, toss, contact and follow through.
3 Points	Student performs good overhand serves 75-90% of the time, executing the 4 important components of a successful serve: foot position, toss, contact and follow through.
2 Points	Student performs good overhand serves 60-75% of the time, executing the 4 important components of a successful serve: foot position, toss, contact and follow through.
1 Point	Student performs good overhand serves less than 60% of the time.

The sample task above shows the format for a performance task along with its item-specific rubric. In this sample, the student is presented with a scenario and instructions for what he or she must do. This is followed by the rubric with the applicable score points.

Criteria for Performance Tasks

1. Performance tasks are worth up to 4 points.
2. Performance tasks should permit multiple solutions to the task presented in the prompt. Rubrics must allow for scoring of multiple solutions.
3. Performance tasks should reference and use materials, tools, and techniques that are appropriate classroom activities for students.
4. Performance tasks must be accompanied by a rubric or scoring guideline that lists descriptions of anticipated responses and the point value for each of the responses.
5. Performance tasks should adhere to all the general guidelines and the constructed response guidelines listed above.

Rubrics

Rubrics, or scoring guides, are used to evaluate constructed response and performance tasks. Rubrics define an ordered range of score points and the criteria associated with the score point. Rubrics may be designed to assess the overall quality of the performance or writing. These rubrics are holistic rubrics, and provide one overall score for the performance or writing. In general, the rubrics on the Health and Physical Education Item bank are holistic rubrics. However, there may be instances where analytic rubrics may be more appropriate. An analytic rubric evaluates multiple aspects of a student's writing or performance, and each aspect has its own set of scoring criteria. The scores for each independent aspect may be combined to result in an overall score.

Example of a holistic rubric:

Stimulus: Explain the physiological benefits of stretching that will help prevent injuries. The response should identify at least three benefits.

Rubric:

- | | |
|----------|--|
| 4 Points | The response correctly and thoroughly assesses the physiological benefits of stretching. The response includes a description of at least three benefits. The response describes the impact of the benefits in preventing injuries. |
| 3 Points | The response assesses the physiological benefits of stretching. The response includes a description of three benefits. The response may have incorrect benefits or inaccurate descriptions. The response may describe the impact of the benefits in preventing injuries. |
| 2 Points | The response may assess the physiological benefits of stretching. The response may include a description of at least two benefits. The response may list incorrect benefits. The response may not describe the impact of the benefits in preventing injuries. |
| 1 Point | The response may not assess the physiological benefits of stretching. The response may not include a description of at least one correct benefit. The response does not describe the impact in preventing injuries. |

Media in Test Items

Test items in the Physical and Health Education Item Bank may include the use of media, through the use of graphics, video clips, or audio clips. Due to the nature of the content area, some benchmarks may be best assessed by utilizing images, video clips, or audio clips. It is important to select images that will render as high-quality in print and on computer displays. Below are the specifications for media that may appear in the stimulus or in response options.

Types

The purpose of the Physical and Health Education items is to measure student achievement in understanding the appropriate way to perform specified activities, the skills necessary to engage in these activities, the proper use of equipment required in certain courses, and the safety measures necessary when performing certain behaviors. There are many real-world applications in Health and Physical Education courses that can best be represented visually, instead of described in words. This can be accomplished in items through use of a graphic, a series of graphics, or through video or audio clips. Images may include black and white line illustrations, color illustrations, or photographs. The use of photographs and video clips will allow for some display of the aspects of the dimensionality and time sequence of a physical activity.

All images should be clear and should demonstrate the focus of the image. The image should be in focus, without excess background material, or anything to distract the eye from the intent of the image.

Images should be at a resolution of 300 dpi and saved as an EPS file. Color images should be saved with the color settings of CMYK (not RGB).

Video clips should be clear and should demonstrate the focus of the question without superfluous background material present. They should not contain more than 2 or 3 people engaged in the activity that is being illustrated. The clip should start at a static position and then show the action to be illustrated clearly from the start to the finish. The clip should last no longer than 1 minute.

Sources

Graphics should be of non-copyrighted images in the public domain or produced or commissioned by the item writer expressly for the Physical and Health Education item bank. Commissioned graphics produced by the item writers for Florida's Health and Physical Education item bank and related products will be the property of the FLDOE. The sources must be cited when the items are turned in for approval.

Characteristics

Graphics must be clear and easy to reproduce as well as authentic. Graphics that require the test taker to have prior or specialized knowledge that is not consistent with the NGSSS should not be included.

Graphics must function as intact pieces. They must also contain the recognizable key concepts which they are intended to illustrate.

Graphics produced for the Florida Physical and Health Education Item Bank should reflect the same qualities of art expressed in the NGSSS. Graphics should present subject matter that is grade-level appropriate.

The selection of public domain graphics must follow the same rigorous review process as all other types of graphics. They should be reviewed by the item writer for any bias and sensitivity issues and grade-level appropriateness.

Content

Graphics should be interesting and appealing to students at the grades for which the graphic is intended. Graphics at a given grade level should include a range of age-appropriate images, drawings, photos, or videos. Also, graphics should contain conceptually appropriate and relevant subjects. Graphics with controversial or offensive content should not be included in the item bank. Confusing or emotionally charged subjects should also be avoided. References to trademarks, commercial products, and brand names should not be included.

Modifications

A public domain graphic or a work of art should not be modified but should be shown as intended by the artist. However, graphics developed specifically for the Florida Physical and Health Education Item Bank that are otherwise appropriate may be modified, through an iterative process with the graphic artist.

Graphic Features

Graphics may include text boxes and other labels, legends, keys and/or captions. Graphics should also reflect multicultural diversity and avoid gender stereotyping.

Item writers must not develop items where the correct response is dependent upon recognition of color, unless required by the benchmark. If a reference to color is used in an item, the color must be labeled with appropriate text. All graphics must be high quality.

Diversity

Graphics should bring a range of cultural diversity to the test. Characters, settings, and situations should reflect the variety of interests and backgrounds that make up Florida's student population. Graphics should represent and or be created by people of different cultures and races; however, graphics about culture- or region-specific graphics should not create an advantage or disadvantage for any particular group of students with a particular characteristic, including gender, race, ethnicity, religion, socioeconomic status, disability, or geographic region. These kinds of graphics must include text with sufficient information to allow a student to answer the question.

Length of Video Clip

Video clips should be a maximum of one minute in length in an effort to minimize test-taking time and keep test takers engaged. Item writers should make their best efforts to ensure that the information required to answer the test question should not require more than one minute of viewing a video.

Item Difficulty and Cognitive Complexity

Often, item difficulty and item complexity are two terms that some people use interchangeably. However, item difficulty is not the same as item complexity.

Item Difficulty

Item difficulty refers to the actual percentage of students who chose the correct answer after an item appears on a test. In general, the items break down into the following categories:

- Easy** More than 70 percent of students are likely to respond correctly.
- Average** Between 40 percent and 70 percent of the students are likely to respond correctly.
- Challenging** Less than 40 percent of the students are likely to respond correctly.

Cognitive Complexity

Cognitive complexity refers to the cognitive demand of the item, or what the student needs to do in his/her head to arrive at the correct answer. The cognitive classification system is based upon Dr. Norman L. Webb's Depth of Knowledge (DOK) levels.¹ The rationale for classifying an item by its DOK level of complexity focuses on the *expectations made by the item*, not on the *ability of the student*. When classifying an item's demands on thinking (i.e., what the item requires the student to recall, understand, analyze, and do), it is assumed that the student is familiar with the basic concepts of the task. Since Health and Physical Education courses are not grade-level specific, the cognitive complexity of an item depends on the benchmark associated with the item.

The categories—low complexity, moderate complexity, and high complexity—form an ordered description of the demands an item may make on a student. For example, Low-complexity items may require a student to solve a one-step problem. Moderate-complexity items may require multiple steps. High-complexity items may require a student to analyze and synthesize information. The distinctions made in item complexity ensure that items will assess the depth of student knowledge at each benchmark. The intent of the item writer weighs heavily in determining the complexity of an item.

The pages that follow illustrate some of the varying demands that items might make at each complexity level for Florida's Physical and Health Education Item Bank. Note that items may fit one or more descriptions. In most instances, these items are classified at the highest level of complexity demanded by the item. Caution must be used in referring to the table on page 23, which describes activities at each cognitive complexity level. This table is provided for ease of reference, but the ultimate determination of the item complexity should be made considering the overall cognitive demand placed on a student.

Item writers are expected to evaluate their items in terms of cognitive complexity and include this on the item template. Items should generally be targeted to the highest level of complexity as appropriate to the assessed benchmark, though some benchmarks call for items at varying levels. When this is the case, writers should take care to cover the range of levels that are appropriate and not create items only at the lower ranges.

¹ Webb, Norman L. and others. "Webb Alignment tool" 24 July 2005. Wisconsin Center of Educational Research. University of Wisconsin-Madison. 2 Feb. 2006. <http://www.wcer.wisc.edu/WAT/index.aspx>.

Low Complexity

Low Complexity items require students to recall, observe, question, or represent basic facts. For a low-complexity item, the student would be expected to demonstrate simple skills or abilities. A low-complexity item requires only basic understanding—often verbatim recall or simple understanding of a course specific term or process.

Below is an example of a low-complexity item.

Course: Basketball

Benchmark: PE.912.M.1.32 Apply sport specific skills in a variety of game settings.

During the 4th quarter of a basketball game, Brittany received a pass at the top of the key. Brittany is now in a triple threat position. What are the three parts of a “triple threat”?

- * A. shoot, pass, or dribble
- B. head fake, dribble, or pass
- C. pass, cut, or rebound
- D. pick, block, or pass

Moderate Complexity

Moderate complexity items require two steps: identification and subsequent processing of that information. Students are expected to make inferences and may encounter items that include words such as classify, organize, and compare. Depending on the objective of a particular moderate level item, students may also be required to explain, describe, or identify cause and effect relationships.

Below is an example of a moderate complexity item.

Course: Volleyball 2

Benchmark: PE.912.C.1.28 Interpret and apply the rules associated with specific course activities.

Jessica’s volleyball class is playing a practice game. Jessica is a back row player on Team A. She crosses in front of the attack line to make her attack, and when the ball is completely above the height of the net, she makes contact. Two players on Team B, come up to block the ball. One of the players who came up to block is a front row player, and the other is a back row player. They both contact the ball on the block. What should the call be?

- * A. It is a double foul and should be replayed.
- B. The procedure is correct and play continues.
- C. Team B’s front row player made an illegal block.
- D. Jessica’s attack was illegal because she’s a back row player.

High Complexity

High complexity items make heavy demands on student thinking. Students may be asked to explain, generalize, or make multiple connections. High complexity items require several steps involving abstract reasoning and planning. Students must be able to support their thinking.

Below is an example of a high complexity item.

Course: HOPE (Variation)

Benchmark: HE.912.C.2.9 Analyze how culture supports and challenges health beliefs, practices, and behaviors.

Stimulus: In some cultures, food and eating is a cornerstone of their lifestyle. Family mealtime is the most important part of the day. Celebrations and religious events often include large, shared meals. Eating well and eating a lot is part of their culture. Write one or two paragraphs explaining how being part of this culture would support and/or challenge one's health beliefs, practices and behaviors.

Rubric:

4 Points	The response provides a thorough analysis of how being part of this culture would support and/or challenge one's health beliefs, practices and behaviors. The response provides a thorough and accurate explanation. The response includes many relevant details to support his/her claims.
3 Points	The response provides an analysis of how being part of this culture would support and/or challenge one's health beliefs, practices and behaviors. The response provides a mostly accurate explanation. The response includes some details and examples to support his/her claims. The response is generally clear, but may include a few inaccuracies.
2 Points	The response provides a partial analysis of how being part of this culture would support and/or challenge one's health beliefs, practices and behaviors. The response provides a partially accurate explanation. The response may include a few details to support his/her claims. The response may include some inaccuracies.
1 Point	The response provides a poor analysis of how being part of this culture would support and/or challenge one's health beliefs, practices and behaviors. The explanation provided is minimal and/or vague. The response may not provide enough information to show evidence of student understanding. The response includes many inaccurate statements.

The following table is provided for ease of reference; however, caution must be used in referring to this table describing activities at each cognitive complexity level. The ultimate determination of an item’s cognitive complexity should be made considering the intent of the overall cognitive demand placed on a student.

Examples of Physical and Health Education Activities Across Cognitive Complexity Levels		
Low Complexity	Moderate Complexity	High Complexity
<ul style="list-style-type: none"> • Define a common term • Complete a one step process • Write a symbol or a list • Label a diagram • Identify types of equipment 	<ul style="list-style-type: none"> • Summarize a procedure • Discuss the importance of a topic • Demonstrate multi-step procedures 	<ul style="list-style-type: none"> • Evaluate situations • Critique the behaviors of a peer • Conduct a research experiment • Analyze a situation and develop a plan of action • Create a cardiorespiratory enhancing workout

An effective test has:

- 15–25% low complexity Items
- 50–60% moderate complexity items
- 15–20% high complexity Items

Content experts from each Physical and Health Education course represented in the Item Bank collaborated to decide upon a test blueprint. The test blueprint generally followed the percentages defined above. Our experts however, decided which benchmarks would be appropriate to assess, and in what capacities. These blueprints will be revisited in year 2 of the project.

Guide to the Item Specifications

What is an item specification?

An item specification is a guideline for item writers and reviewers. Item specifications define the context and content which belongs in a test item. Item specifications link content standards to test items.

It is important to note the organization of the content standards and benchmarks.

Physical and Health Education Benchmark Classification System

Each benchmark in the NGSSS is coded with a system of numbers and letters.

- The two letters in the *first position* of the code identify the **Subject Area** (e.g., PE for Physical Education, HE for Health Education).
- The number in the *second position* (first number) represents the **Grade Level**. (e.g., 912 for High School)
- The letter in the *third position* represents the **Strand**, or category of knowledge, to which the benchmark belongs.

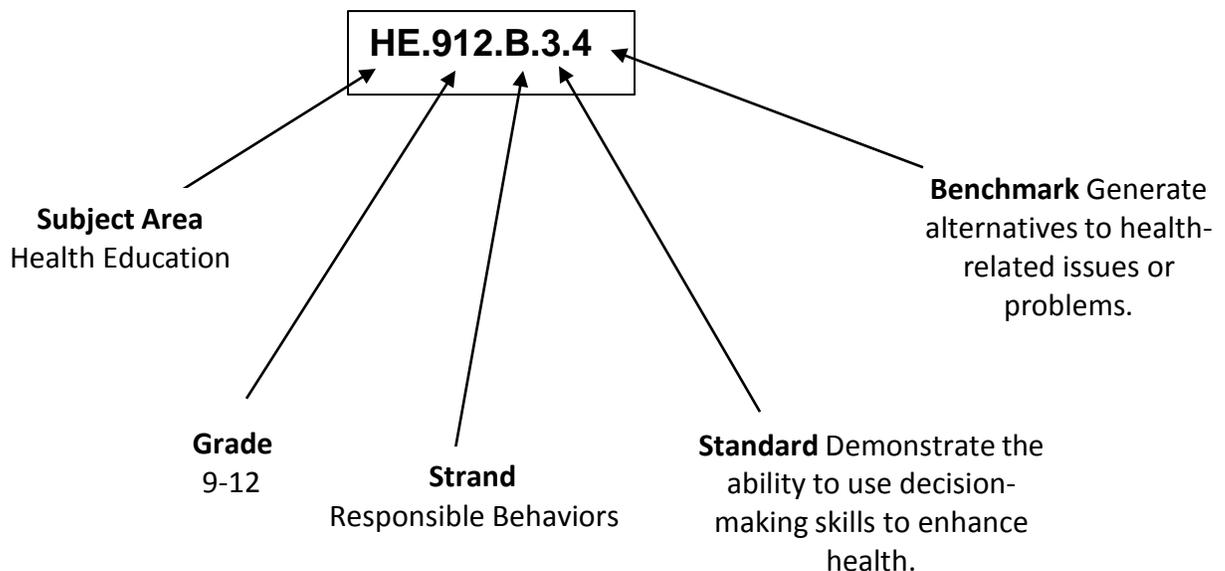
In Physical Education, there are five strands.

- Strand 1: Movement Competency
- Strand 2: Cognitive Abilities
- Strand 3: Lifetime Fitness
- Strand 4: Responsible Behaviors and Values
- Strand 5: Innovation, Technology, and the Future

In Health Education, there are three strands:

- Strand 1: Responsible Behavior
- Strand 2: Concepts
- Strand 3: Promotion

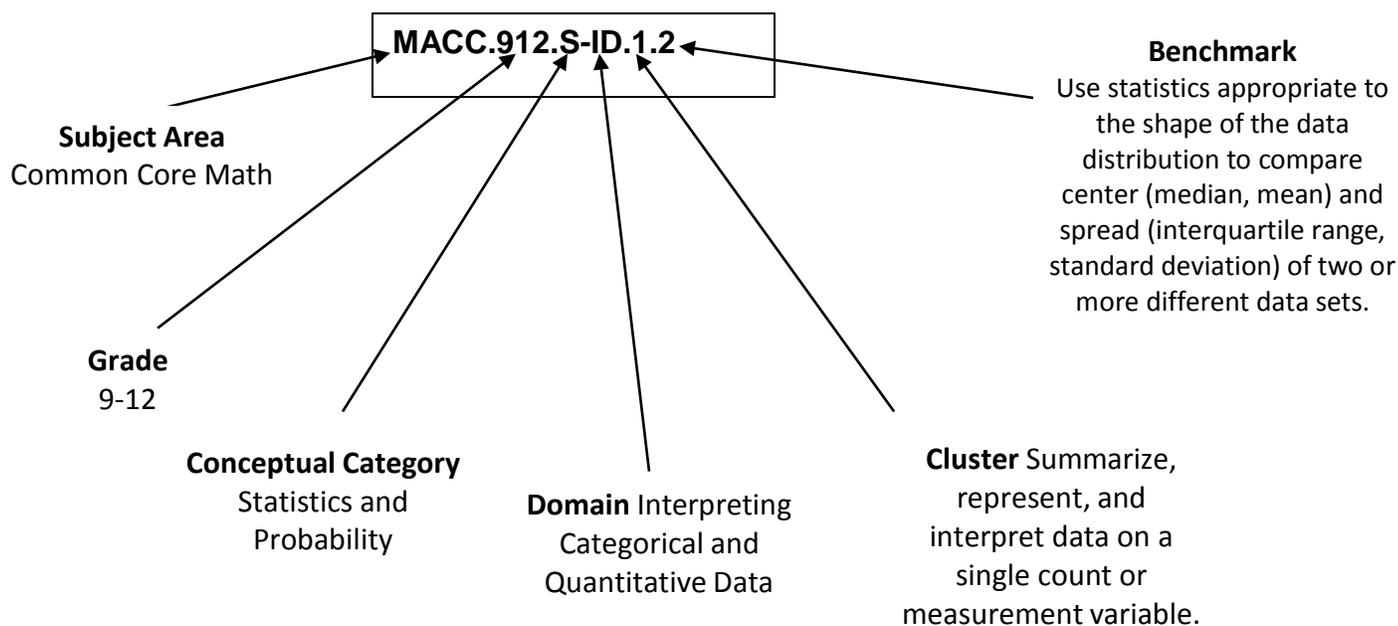
- The number in the *fourth position* of the code represents the **Standard** for the benchmark.
- The number in the *fifth position* shows the specific **Benchmark** that falls under the specified strand and within the standard.



Common Core Benchmark Classification System -Math

Each benchmark in the CCSS is coded with a system of numbers and letters.

- The four letters in the *first position* of the code identify the **Subject Area** (e.g., MACC).
- The number in the *second position* (first number) represents the **Grade Level**. (e.g., 910 9th and 10th Grade.)
- The letter(s) in the *third position* represents the **Conceptual Category**.
In High School Math, there are 6 conceptual categories.
 - Strand 1: Number and Quantity
 - Strand 2: Algebra
 - Strand 3: Functions
 - Strand 4: Modeling
 - Strand 5: Geometry
 - Strand 6: Statistics and Probability
- The letter(s) in the *fourth position* represents the **Domain**, as it relates to the conceptual category. (e.g., Interpreting Categorical and Quantitative Data)
- The number in the *fifth position* of the code represents the **Cluster** for the benchmark.
- The number in the *sixth position* shows the specific **Benchmark** that falls under the specified strand and within the standard.



Item Specifications Template and Definitions

Item specifications for the Physical and Health Education Assessment bank include the following components:

Standard: Each standard is a general statement of expected student achievement within a course in the NGSSS.

Benchmark: Benchmarks are specific statements of expected student achievement under course-specific Standards.

Depth of Knowledge: Based on Webb's Depth of Knowledge, items may be Low Complexity, Moderate Complexity, or High Complexity. Item specifications may include more than one category.

Item Types: This section of the item specification identifies what types of items may be written to address the benchmark. Items may be Multiple Choice, Constructed Response, or Performance Task.

Content Limits: The content limits define the scope of content knowledge that will be assessed (e.g., specific elements that can be compared or contrasted) and in some cases, indicate areas of the benchmark that will not be assessed. For some benchmarks, additional information is provided to clarify specific directions in developing test items.

Stimulus Attributes: Stimulus attributes define the types of stimuli that will be used in the development of items, including appropriate context or content suitable for assessing the particular benchmark.

Response Attributes: The response attributes define the characteristics of the responses, either which the student will produce, or those that appear in the multiple choice item.

Sample Item: Sample items for each benchmark are provided. Item specifications (at this time) may include one or more sample item. The sample item is one of the listed item types. The sample items are presented in a format similar to the test, and the correct answer or scoring rubric for each sample item is provided.

Guidelines for Item Writers

Item writers participating in the project must be certified to teach the course for which they are writing items. The teachers must have a comprehensive knowledge of the content covered by their course curriculum, as well as an intimate understanding of the NGSSS benchmarks for the Health and Physical Education courses they teach. Additionally, since item writers are familiar with the age group of the students who are the audience for this item bank, item writers should also be familiar with the cognitive abilities of these students. Item writers should use their best judgment in writing items that measure Health and Physical Education benchmarks of the NGSSS without introducing extraneous elements that may interfere with the test's validity. Item writers should know and respect the guidelines established in this document as well as appreciate the spirit of developing test content that allows students to perform at their best.

Item writers for Florida's Health and Physical Education Item Bank must submit items in an approved Item Writing Template which includes the following information about each item.

Format	Item writers must submit items in the template provided in Appendix B. All appropriate sections of the template should be completed before the items are submitted.
Sources	Item writers are expected to provide sources for all images included in the item. Acceptable sources for art work are public domain sources that do not require copyright fees for use. Photos of student images and videos must be accompanied by a completed permission form.
Correct Response	<p>Multiple-Choice Items: Item writers must indicate which option is the correct answer.</p> <p>Constructed Response Items: Item writers must provide a scoring rubric that includes the required ideas/wording that must be present for each score point.</p> <p>Performance Tasks: Item writers must provide a scoring rubric that clearly delineates the activities or actions that must be present for each score point.</p>
Option Rationales	<p>Multiple-Choice Items: All response options for a multiple choice item must be accompanied by a rationale. The rationales defend the plausibility of the response option.</p> <p>For the correct option the rationale must state why the option is correct based on the benchmark being assessed.</p> <p>For incorrect options the rationales must state why the option is plausible and why it is incorrect based on the benchmark being assessed.</p>
Cognitive Complexity	Item writers are expected to write items that are low, moderate, or high complexity. The cognitive complexity of the item must be one indicated on the item specification. Item writers should take care to cover the range of levels that are appropriate and to not create items only at the lower ranges.

Submission of Items	<p>When submitting items, item writers must balance several factors. Item submissions should</p> <ul style="list-style-type: none"> • include items for the benchmarks and cognitive complexities assigned to the item writer; • have an approximate balance of the correct response between the answer choices for multiple-choice items; • have an equal balance of male and female names and include names representing different ethnic groups in Florida; • have an equal balance of male and female students at the appropriate grade level in the stimulus and/or option photographs and drawings; and • have an equal balance of ethnic groups in Florida represented in the stimulus and/or option photographs and drawings.
Electronic Submission	<p>Items will be submitted into the online collaboration tool, Basecamp. Project participants will be invited to create a password-protected Basecamp account. Information about how to upload items into the system will be provided to participants. Items will then be input into the AIR temporary item banking system.</p>

Review Procedures

Before being accepted into Florida’s Health and Physical Education Item Bank, all items must pass several levels of review. Florida educators, in conjunction with the DOE and the LEAs, scrutinize all items prior to accepting them for placement in the item bank. The graphics and items are reviewed for content characteristics, potential bias, and any issues of concern to Florida stakeholders. Concerns expressed during the reviews must be resolved satisfactorily before the graphics and items are placed in the item bank.

Review of Test Items

The DOE, the LEAs, and a committee made up of select Florida educators with experience and expertise in Health and Physical Education instruction review all test items during the item development process. The content specialists at the DOE review and edit items, judging them for overall quality and suitability for the tested grade level.

Prior to submission to the DOE for their review, items are reviewed by project staff, which includes Health and Physical Education teachers from across the state who have agreed to participate in the Review stage of the development process.

The reviewers will be focusing on the content of the items, issues of bias and sensitivity, and style and formatting issues. The primary focus however, is the content validity, whether or not the item is a valid measure of the designated NGSSS benchmark, and if the item aligns with the item specifications.

Item writers and reviewers will be provided a checklist to guide their thinking as it relates to their specific task. The Item Writing and Reviewing Checklist is found in Appendix C. The checklist directs the reviewers to examine the item as a whole, and also to independently examine the stimulus and question stem as well as the response options and answer key. The Checklist is separated into sections for each of those components and includes guiding questions. Item Reviewers will use the Checklist with every item they review.

Feedback on the item will be recorded on an Item Review Cover Sheet for each item. The Item Review Cover Sheet, found in Appendix D, correlates with the Checklist. However, the coversheet provides only a few key questions that relate to the various aspects of the item, instead of the exhaustive list that is found on the Checklist. Reviewers are directed to mark which aspects of the item they have concerns about, and provide specific written feedback on the Cover Sheet. Then, the Reviewer will rate the item based on a rubric provided on the Cover Sheet. The reviewers' feedback is given to the original item writer, and the item writer has time to implement the feedback and make the necessary changes. At that point, the item may be submitted into the Item Banking System.

Review for Potential Bias and Sensitivity Issues

The Department of Education provided documentation to guide item reviews for bias and sensitivity. We incorporated this document into our review checklist and provided the FLDOE documentation to all of our item writers and reviewers. Both the graphics and items are reviewed by Florida educators for the following kinds of bias: gender, racial, ethnic, linguistic, religious, geographic, and socioeconomic. Reviews also include consideration of issues relevant to individuals with disabilities.

The purpose for the sensitivity review is to ensure that the topic or information presented in the item does not cause any strong student reaction which would negatively affect the student's performance on the item. Topics that may be offensive, or that students may have particularly strong negative associations or experiences with should be avoided. The topic of the item should not distract from the purpose of the assessment item. Some topics that may be deemed inappropriate to include on assessment items are wildfires, hurricanes, religion, death, or child abuse.

Appendix A: Physical and Health Education Standards and Benchmarks

Each Physical and Health Education Course has a set of course standards which outlines what the students should be learning throughout the course. The standards are organized into clusters, which relate to the content addressed by each of the standards. Each standard has associated benchmarks, which are more specific expressions of the standard. The teacher is required to teach all of the benchmarks in the course. The item bank will include items for each assessable benchmark for the courses included in this grant.

Physical Education Standards and Benchmarks

Strand	Cognitive Abilities
Standard 1	Identify, analyze, and evaluate movement concepts, mechanical principles, safety considerations, and strategies/tactics regarding movement
PE.912.C.1.1	Identify and describe the critical elements of a basic water rescue.
PE.912.C.1.2	Understand and apply terminology and etiquette in dance.
PE.912.C.1.3	Analyze through observation the movement performance of self and others.
PE.912.C.1.4	Choreograph complex dance sequences alone, with a partner, or in a small group.
PE.912.C.1.5	Analyze the relationship between music and dance.
PE.912.C.1.6	Compare and contrast the health-related benefits of various physical activities.
PE.912.C.1.7	Evaluate the effectiveness of specific warm-up and cool-down activities.
PE.912.C.1.8	Differentiate between the three different types of heat illnesses associated with fluid loss.
PE.912.C.1.9	Explain the precautions to be taken when exercising in extreme weather and/or environmental conditions.
PE.912.C.1.10	Analyze long-term benefits of participating in regular physical activity.
PE.912.C.1.11	Explain how each of the health-related fitness components (cardiorespiratory endurance, muscular strength, muscular endurance, flexibility, body composition) are improved through the application of training principles.
E.912.C.1.12	Compare and contrast aerobic versus anaerobic activities.
PE.912.C.1.13	Document food intake, calories consumed, and energy expended through physical activity and analyze the results.
PE.912.C.1.14	Compare and contrast the skill-related components of fitness (speed, coordination, balance, power, agility, reaction time) used in various physical activities.
PE.912.C.1.15	Calculate individual target heart rate zone and analyze how to adjust intensity level to stay within the desired range.
PE.912.C.1.16	Explain the methods of monitoring levels of intensity during aerobic activity.
PE.912.C.1.17	Assess physiological effects of exercise during and after physical activity.
PE.912.C.1.18	Differentiate between fact and fallacy as it relates to consumer physical fitness products and programs.
PE.912.C.1.19	Choreograph complex sequences alone, with a partner, or in a small group.
PE.912.C.1.20	Know various ways in which physical conflict can be resolved appropriately.
PE.912.C.1.21	Diagram, explain, and justify the use of advanced offensive, defensive, and transition strategies and tactics.
PE.912.C.1.22	Explain the skill-related components of balance, reaction time, agility, coordination, power, and speed and how they enhance performance levels.

PE.912.C.1.23	Apply appropriate technology and analyze data to evaluate, monitor, and/or improve performance.
PE.912.C.1.24	Analyze the mechanical principles as they apply to specific course activities.
PE.912.C.1.25	Analyze and evaluate the risks, safety procedures, rules, and equipment associated with specific course activities.
PE.912.C.1.26	Evaluate skill patterns of self and/or partner by detecting and correcting mechanical errors.
PE.912.C.1.27	Compare and contrast how movement skills from one physical activity can be transferred and used in other physical activities.
PE.912.C.1.28	Interpret and apply the rules associated with specific course activities.
Strand	Lifetime Fitness
Standard 1	Participate regularly in physical activity.
PE.912.L.1.1	Participate in a variety of physical activities to meet the recommended number of minutes of moderate to vigorous physical activity (MVPA) beyond physical education on five or more days of the week.
PE.912.L.1.2	Participate in a variety of activities that promote cardiorespiratory fitness, muscular strength and endurance, flexibility, and body composition.
PE.912.L.1.3	Participate in a variety of activities that promote effective stress management.
PE.912.L.1.4	Utilize the in-school and community opportunities for participation in a variety of physical activities.
PE.912.L.1.5	Participate regularly in health-enhancing activities outside the physical education class setting.
PE.912.L.1.6	Utilize knowledge of the risks and safety factors that may affect physical activity throughout life.
Standard 2	Develop and implement a personal fitness program to achieve and maintain a health-enhancing level of physical fitness.
PE.912.L.2.1	Demonstrate achievement and maintenance of a health-enhancing level of personal fitness by designing, implementing, self-assessing, and modifying a personal fitness program.
PE.912.L.2.2	Demonstrate program planning skills by setting goals, devising strategies, and making timelines for a personal fitness program.
PE.912.L.2.3	Use a variety of resources including available technology to assess, design, and evaluate their personal physical activity plan.
PE.912.L.2.4	Apply the principles of training and conditioning in accordance with personal goals.
PE.912.L.2.5	Assess and evaluate the use of a variety of physical activities in developing a personal fitness program.
PE.912.L.2.6	Analyze health-related problems associated with inadequate levels of cardiorespiratory endurance, muscular strength and endurance, flexibility, and body composition.
PE.912.L.2.7	Evaluate how to make changes in an individual wellness plan as lifestyle changes occur.

Strand	Movement Competency:
Standard 1	Demonstrate competency in many and proficiency in a few movement forms from a variety of categories (aquatics, dance, extreme sports, fitness education, gymnastics, individual/dual sports, outdoor pursuits, self-defense, team sports).
PE.912.M.1.1	Demonstrate critical elements of basic skills relating to aquatics.
PE.912.M.1.2	Demonstrate proficiency in combination of motor skills related to aquatics.
PE.912.M.1.3	Perform a basic water rescue, with or without equipment, without entering the water.
PE.912.M.1.4	Perform refinement of one or more swim strokes to enhance efficiency, power, and cardiorespiratory endurance in a variety of aquatics settings.
PE.912.M.1.5	Apply strategies for self-improvement based on individual strengths and needs.
PE.912.M.1.6	Select appropriate music for dance forms and choreograph dance movements to music.
PE.912.M.1.7	Perform advanced dance sequences from a variety of dances accurately and with correct technique.
PE.912.M.1.8	Design and perform a creative movement sequence while working with a small or large group, with or without equipment/props.
PE.912.M.1.9	Demonstrate complex skills and advanced rhythmic movements in dance.
PE.912.M.1.10	Apply sport specific skills in simulation and in real-life applications.
PE.912.M.1.11	Demonstrate competency in two or more extreme sports activities.
PE.912.M.1.12	Select and perform complex movements using a variety of equipment which lead to improved or maintained muscular strength and endurance.
PE.912.M.1.13	Perform a student designed cardiorespiratory enhancing workout.
PE.912.M.1.14	Utilize selected technology to assess, enhance, and maintain health and skill-related fitness levels.
PE.912.M.1.15	Select and apply sports/activity specific warm-up and cool-down techniques.
PE.912.M.1.16	Apply the principles of training and conditioning to accommodate individual needs and strengths.
PE.912.M.1.17	Demonstrate basic cardiopulmonary resuscitation (CPR) procedures.
PE.912.M.1.18	Demonstrate a variety of gymnastics skills with a level of control.
PE.912.M.1.19	Use correct body alignment, strength, flexibility, and coordination in the performance of technical movements.
PE.912.M.1.20	Perform complex combinations and sequences demonstrating smooth transitions while alone, with a partner, or in a small group.
PE.912.M.1.21	Demonstrate the relationship between complex dance elements and rhythmic movements related to educational gymnastics skills and sequences.
PE.912.M.1.22	Demonstrate proficiency in advanced combinations of motor skills for a variety of individual and dual sports.
PE.912.M.1.23	Demonstrate proficiency of critical elements when striking with an object/implement.
PE.912.M.1.24	Apply a combination of complex movement patterns in a game setting.
PE.912.M.1.25	Apply the appropriate speed and generation of force when running sprints or distance, throwing, jumping, and striking.
PE.912.M.1.26	Analyze and apply offensive, defensive, and transition strategies and tactics to reflect a higher order of thinking.
PE.912.M.1.27	Demonstrate proficiency in a variety of outdoor pursuit activities.
PE.912.M.1.28	Apply strategies and tactics in a variety of outdoor pursuits.
PE.912.M.1.29	Demonstrate proficiency in self-defense movement skills.

PE.912.M.1.30	Combine and apply movement patterns from simple to complex.
PE.912.M.1.31	Demonstrate advanced offensive, defensive, and transition strategies and tactics.
PE.912.M.1.32	Apply sport specific skills in a variety of game settings.
PE.912.M.1.33	Practice complex motor activities in order to improve performance.
PE.912.M.1.34	Demonstrate use of the mechanical principles as they apply to specific course activities.
PE.912.M.1.35	Select proper equipment and apply all appropriate safety procedures necessary for participation.
Strand	Responsible Behaviors
Standard 1	Exhibit responsible personal and social behavior that respects self and others in physical activity settings.
PE.912.R.1.1	Act independently of peer pressure both in and out of school.
PE.912.R.1.2	Develop strategies for including persons of diverse backgrounds and abilities while participating in a variety of physical activities.
PE.912.R.1.3	Demonstrate responsible behaviors during physical activities.
PE.912.R.1.4	Maintain appropriate personal, social, and ethical behavior while participating in a variety of physical activities.
PE.912.R.1.5	Demonstrate appropriate etiquette, care of equipment, respect for facilities, and safe behaviors while participating in a variety of physical activities.
Standard 2	Value physical activity for health, enjoyment, challenge, self-expression, and/or social interaction.
PE.912.R.2.1	Select and participate in a variety of physical activities outside of the school setting that contribute to personal enjoyment and the attainment or maintenance of a healthy lifestyle.
PE.912.R.2.2	Discuss physical activities from which benefits can be derived.
PE.912.R.2.3	Explore the role of games, sports, and/or physical activities in other cultures.

Health Education Standards and Benchmarks

Strand	RESPONSIBLE BEHAVIOR
Standard 1	Demonstrate the ability to access valid health information, products, and services to enhance health.
HE.912.B.1.1	Verify the validity of health information, products, and services.
HE.912.B.1.2	Compile data reflecting the accessibility of resources from home, school, and community that provide valid health information.
HE.912.B.1.3	Evaluate the accessibility of products and services that enhance health.
HE.912.B.1.4	Justify when professional health services or providers may be required.
HE.912.B.1.5	Critique valid and reliable health products and services.
HE.912.B.1.6	Justify the validity of a variety of technologies to gather health information.
Standard 2	Demonstrate the ability to use interpersonal communication skills to enhance health and avoid or reduce health risks.
HE.912.B.2.1	Explain skills needed to communicate effectively with family, peers, and others to enhance health.
HE.912.B.2.2	Assess refusal, negotiation, and collaboration skills to enhance health and avoid or reduce health risks.
HE.912.B.2.3	Demonstrate strategies to prevent, manage, or resolve interpersonal conflicts without harming self or others.
HE.912.B.2.4	Analyze the validity of ways to ask for and offer assistance to enhance the health of self and others.
Standard 3	Demonstrate the ability to use decision-making skills to enhance health.
HE.912.B.3.1	Determine the value of applying a thoughtful decision-making process in health-related situations.
HE.912.B.3.2	Examine barriers that can hinder healthy decision-making.
HE.912.B.3.3	Assess whether individual or collaborative decision-making is needed to make a healthy decision.
HE.912.B.3.4	Generate alternatives to health-related issues or problems.
HE.912.B.3.5	Appraise the potential short-term and long-term outcomes of each alternative on self and others.
HE.912.B.3.6	Employ the healthiest choice when considering all factors in making a decision.
Standard 4	Demonstrate the ability to use goal-setting skills to enhance health.
HE.912.B.4.1	Evaluate personal health practices and overall health status to include all dimensions of health.
HE.912.B.4.2	Formulate a plan to attain a personal health goal that addresses strengths, needs, and risks.
HE.912.B.4.3	Implement strategies and monitor progress in achieving a personal health goal.
HE.912.B.4.4	Formulate an effective long-term personal health plan.

Strand	CONCEPTS
Standard 1	Formulate an effective long-term personal health plan. Comprehend concepts related to health promotion and disease prevention to enhance health.
HE.912.C.1.1	Predict how healthy behaviors can affect health status.
HE.912.C.1.2	Interpret the interrelationships of mental/emotional, intellectual, physical, and social health.
HE.912.C.1.3	Evaluate how environment and personal health are interrelated.
HE.912.C.1.4	Analyze how heredity and family history can impact personal health.
HE.912.C.1.5	Propose strategies to reduce or prevent injuries and health problems.
HE.912.C.1.6	Evaluate the relationship between access to health care and health status.
HE.912.C.1.7	Assess the degree of susceptibility to injury, illness or death if engaging in unhealthy/risky behaviors.
HE.912.C.1.8	Analyze strategies for prevention, detection, and treatment of communicable and chronic diseases.
Standard 2	Analyze the influence of family, peers, culture, media, technology, and other factors on health behaviors.
HE.912.C.2.1	Analyze how the family influences the health of individuals.
HE.912.C.2.2	Compare how peers influence healthy and unhealthy behaviors.
HE.912.C.2.3	Assess how the school and community can affect personal health practice and behaviors.
HE.912.C.2.4	Evaluate how public health policies and government regulations can influence health promotion and disease prevention.
HE.912.C.2.5	Evaluate the effect of media on personal and family health.
HE.912.C.2.6	Evaluate the impact of technology on personal, family, and community health.
HE.912.C.2.7	Assess the consequences of health risk behaviors.
HE.912.C.2.8	Analyze how the perceptions of norms influence healthy and unhealthy behaviors.
HE.912.C.2.9	Analyze how culture supports and challenges health beliefs, practices, and behaviors.
Strand	PROMOTION
Standard 1	Demonstrate the ability to practice advocacy, health-enhancing behaviors, and avoidance or reduction of health risks for oneself.
HE.912.P.1.1	Analyze the role of individual responsibility in enhancing health.
HE.912.P.1.2	Demonstrate a variety of healthy practices and behaviors that will maintain or improve health.
HE.912.P.1.3	Critique a variety of behaviors that avoid or reduce health risks.
Standard 2	Demonstrate the ability to advocate for individual, peer, school, family, and community health.
HE.912.P.2.1	Utilize current, accurate data/information to formulate a health-enhancing message.
HE.912.P.2.2	Demonstrate how to influence and support others in making positive health choices.
HE.912.P.2.3	Work cooperatively as an advocate for improving personal, family and community health.
HE.912.P.2.4	Adapt health messages and communication techniques to a specific target audience.

Appendix B: Item Writing Template

Project: B- Physical and Health Education

Course Name:

Item Writer ID:

Webb's Cognitive Level:

Benchmark/standard:

Item:

Choices:

Correct
Answer

A)	
B)	
C)	
D)	

For each incorrect choice describe why the answer is plausible.

A)	
B)	
C)	
D)	

Appendix C: Item Writing and Review Checklist

For each test item you submit, review this checklist² alongside of each item.

Overarching questions
<ol style="list-style-type: none">1) Does the item match the benchmark?2) Does the item meet the intended complexity?3) Does the item match the item specification in terms of Content Limits, Stimulus Attributes, and Response Attributes?4) Is the item different from the sample item provided in the item specification?5) Is the item free of bias/sensitivity?6) Is the item laid out in a clear and consistent manner?7) Are the grammar, punctuation, and spelling correct?
Reviewing the Stimulus and Question Stem
<p>Look at the stimulus (excluding the question stem), ask yourself:</p> <ul style="list-style-type: none">• If you are including a sample technical text or included a passage,<ul style="list-style-type: none">○ Have you verified that the passage is original, or have you cited all sources for a passage which has been adapted from an existing text?○ Is it a realistic representation of something a student in this course would read?○ Does it relate to the benchmark?○ Is it clear and well written?○ Is all the necessary content present?• If you are including a scenario,<ul style="list-style-type: none">○ Is it realistic?○ Is it grade level appropriate?○ Does it relate to the course?○ Does it relate to the benchmark?○ Are the text and amount of reading required appropriate and reasonable for the grade level (assume 2 grade levels below)?• If you are including data or charts/graphs,<ul style="list-style-type: none">○ Is the data related to the course?○ Is the data realistic and accurate?○ Is the data labeled appropriately?○ In your graph, are the axes labeled and does the graph have a title?• If you are including an image,<ul style="list-style-type: none">○ Is the source cited (if potential copyright issue)?○ Is the image of good quality?○ Is it free of bias/stereotyping? <p>Look specifically at the question stem:</p> <ul style="list-style-type: none">• Is the question stem appropriate for the benchmark?• Does the question stem relate to the passage, scenario, data, chart, graphic, etc. provided?• Is the intent of the question stem clear? Does it <i>clearly</i> ask the student what needs to be done?

² Adapted from:

Reading Item Checklist, CAL/ WIDA Consortium

Haladyna, T. (1999). *Developing and Validating Multiple Choice Test Items*. Mahwah, NJ: Lawrence Erlbaum Associates.

Checklist for Multiple Choice Item Review, USF

- Does the question stem read naturally? Is the question stem concise and well written? Is it free of colloquial language or slang?
- Can the question be answered in a straightforward way without viewing the response options? (e.g., when you cover the response options and read it, can you answer the question? The answer in MOST cases should be YES.)
- Does the natural response to the task statement/question provide evidence that the student can do what is indicated in the benchmark?
- Does the item stem include all the information necessary to understand the question?
- Are repeated terms included in the stem? (e.g., if the response options **all** start “He should ...”; consider including that in the question stem.)
- Does the item avoid logical deduction or “testwiseness”? (e.g.; a word in the question stem shouldn’t be repeated in the key)
- As needed, is there a reminder to seek the answer from the passage (e.g., according to this passage...)?
 - This is important when other background knowledge might lead the student somewhere else. “Based on what you’ve read...” serves to focus the student on the immediate passage.
- Is the stem stated in a positive way (unless it cannot be avoided)?
- If the stem contains the words: **MUST, ONLY, NOT**, etc., are they **bolded** and CAPITALIZED?

Reviewing the Response Options

- Do all of the response options answer the question? (e.g., if the question is “Which activity should she choose?” a response option should not be “happy.” “Happy” is not an activity.)
- Are the response options complete sentences?
 - If so, do they begin with a capital letter, and end with a period?
 - If not, the first letter should NOT be capitalized, and there should NOT be a period.
- Are the length, word frequency, and other surface/deep characteristics of each distractor about equal to each other?
 - If the response options must be different lengths, are they arranged in order from shortest to longest?
- Are the response options arranged in chronological (if applicable) or alphabetical order?
- Are there four response options?
- Are the following absent from all response options:
 - Opposites (e.g., one cancels out the other)
 - Overlapping choices (e.g., options that essentially say the same thing)
 - Unintentionally tricky or “garden path” distractors
 - Impossible or nonsensical distractors
 - Responses that lead the student to the correct answer
 - “All of the above.” or “None of the above.”

For the **Answer Key**, ask yourself:

- Is it a natural answer to the question stem?
- Does it give evidence that the student can do what is indicated in the benchmark?
- Is it clearly the **ONLY** correct answer?
- If you were to cover up all response options and answer the question, would your answer be the key? (must be YES to avoid trick questions)

For the **Distractors**, ask yourself, “if a student understood the question but does not know the answer”:

- Is each distractor clearly a plausible answer to the task statement/question?
- Is the distractor plausibility section complete in the template?
- Is each distractor clearly incorrect?

Appendix D: Item Review Cover Sheet

Course Name:

Benchmark Number:

Item Writer ID (from the filename):

Item ID (from the filename):

Reviewer ID (This is YOUR number):

Criteria	Yes/No	Comments*
Overarching Questions		
1) Does the item match the benchmark?		
2) Does the item meet the intended complexity?		
3) Does the item match the item specification in terms of Content Limits, Stimulus Attributes, and Response Attributes ?		
4) Is the item different from the sample item provided in the item specification?		
5) Is the item free of bias/sensitivity?		
Stimulus and Question Stem		
6) Are there any questions/concerns with the Stimulus ?		
7) Are there any questions/concerns with the Question Stem ?		
Response Options		
8) Are there any questions/concerns with answer key ?		
9) Are there any questions/concerns with the distractors ?		
Final Check		
10) Is the item laid out in a clear and consistent manner?		
11) Are the grammar, punctuation, and spelling correct?		

* You must complete the "Comments" section if there is a concern with the item.

General comments/feedback on item:

Rating Rubric:

- 1 point The item needs **extensive** revision. There are major problems with the content of this item. It needs an entire rewrite. I have marked errors in the “Overarching Questions” section of the coversheet. I may also have marked errors in the “Stimulus and Question Stem” and/or the “Response Options” sections of the coversheet.
- 2 points The item needs **some** revisions. The item has problems, but can be salvaged using the comments and suggestions I (the reviewer) have provided. I have marked errors in the “Stimulus and Question Stem” and/or the “Response Options” sections of the coversheet.
- 3 points The item needs **minor** revisions. There are only a few small grammar/punctuation/format issues that should be corrected. I have indicated these errors in the “Final Check” section of the coversheet.
- 4 points The item is good and needs no revisions. It is ready to be submitted to DOE.

I’m rating this item: _____