

CODE OF FAIR TESTING PRACTICES IN EDUCATION

Prepared by the Joint Committee on Testing Practices

The Code of Fair Testing Practices in Education (**Code**) is a guide for professionals in fulfilling their obligation to provide and use tests that are fair to all test takers regardless of age, gender, disability, race, ethnicity, national origin, religion, sexual orientation, linguistic background, or other personal characteristics. Fairness is a primary consideration in all aspects of testing. Careful standardization of tests and administration conditions helps to ensure that all test takers are given a comparable opportunity to demonstrate what they know and how they can perform in the area being tested. Fairness implies that every test taker has the opportunity to prepare for the test and is informed about the general nature and content of the test, as appropriate to the purpose of the test. Fairness also extends to the accurate reporting of individual and group test results. Fairness is not an isolated concept, but must be considered in all aspects of the testing process.

The **Code** applies broadly to testing in education (admissions, educational assessment, educational diagnosis, and student placement) regardless of the mode of presentation, so it is relevant to conventional paper-and-pencil tests, computer based tests, and performance tests. It is not designed to cover employment testing, licensure or certification testing, or other types of testing outside the field of education. The **Code** is directed primarily at professionally developed tests used in formally administered testing programs. Although the **Code** is not intended to cover tests made by teachers for use in their own classrooms, teachers are encouraged to use the guidelines to help improve their testing practices.

The **Code** addresses the roles of test developers and test users separately. Test developers are people and organizations that construct tests, as well as those that set policies for testing programs. Test users are people and agencies that select tests, administer tests, commission test development services, or make decisions on the basis of test scores. Test developer and test user roles may overlap, for example, when a state or local education agency commissions test development services, sets policies that control the test development process, and makes decisions on the basis of the test scores.

Many of the statements in the **Code** refer to the selection and use of existing tests. When a new test is developed, when an existing test is modified, or when the administration of a test is modified, the **Code** is intended to provide guidance for this process.

1

1 The Code is not intended to be mandatory, exhaustive, or definitive, and may not be applicable to every situation. Instead, the Code is intended to be aspirational, and is not intended to take precedence over the judgment of those who have competence in the subjects addressed.

The **Code** provides guidance separately for test developers and test users in four critical areas:

- A. Developing and Selecting Appropriate Tests
- B. Administering and Scoring Tests
- C. Reporting and Interpreting Test Results
- D. Informing Test Takers

The **Code** is intended to be consistent with the relevant parts of the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], and National Council on Measurement in Education [NCME], 1999). The **Code** is not meant to add new principles over and above those in the *Standards* or to change their meaning. Rather, the **Code** is intended to represent the spirit of selected portions of the *Standards* in a way that is relevant and meaningful to developers and users of tests, as well as to test takers and/or their parents or guardians. States, districts, schools, organizations and individual professionals are encouraged to commit themselves to fairness in testing and safeguarding the rights of test takers. The Code is intended to assist in carrying out such commitments.

The **Code** has been prepared by the Joint Committee on Testing Practices, a cooperative effort among several professional organizations. The aim of the Joint Committee is to act, in the public interest, to advance the quality of testing practices. Members of the Joint Committee include the American Counseling Association (ACA), the American Educational Research Association (AERA), the American Psychological Association (APA), the American Speech-Language-Hearing Association (ASHA), the National Association of School Psychologists (NASP), the National Association of Test Directors (NATD), and the National Council on Measurement in Education (NCME).

Copyright 2004 by the Joint Committee on Testing Practices. This material may be reproduced in whole or in part without fees or permission, provided that acknowledgment is made to the Joint Committee on Testing Practices. Reproduction and dissemination of this document are encouraged. This edition replaces the first edition of the **Code**, which was published in 1988. Please cite this document as follows: *Code of Fair Testing Practices in Education*. (2004). Washington, DC: Joint Committee on Testing Practices. (Mailing Address: Joint Committee on Testing Practices, Science Directorate, American Psychological Association, 750 First Street, NE, Washington, DC 20002-4242; <http://www.apa.org/science/jctpweb.html>.) Contact APA for additional copies.

A. Developing and Selecting Appropriate Tests

TEST DEVELOPERS	TEST USERS
Test developers should provide the information and supporting evidence that test users need to select appropriate tests.	Test users should select tests that meet the intended purpose and that are appropriate for the intended test takers.
A-1. Provide evidence of what the test measures, the recommended uses, the intended test takers, and the strengths and limitations of the test, including the level of precision of the test scores.	A-1. Define the purpose for testing, the content and skills to be tested, and the intended test takers. Select and use the most appropriate test based on a thorough review of available information.
A-2. Describe how the content and skills to be tested were selected and how the tests were developed.	A-2. Review and select tests based on the appropriateness of test content, skills tested, and content coverage for the intended purpose of testing.
A-3. Communicate information about a test's characteristics at a level of detail appropriate to the intended test users.	A-3. Review materials provided by test developers and select tests for which clear, accurate, and complete information is provided.
A-4. Provide guidance on the levels of skills, knowledge, and training necessary for appropriate review, selection, and administration of tests.	A-4. Select tests through a process that includes persons with appropriate knowledge, skills, and training.
A-5. Provide evidence that the technical quality, including reliability and validity, of the test meets its intended purposes.	A-5. Evaluate evidence of the technical quality of the test provided by the test developer and any independent reviewers.
A-6. Provide to qualified test users representative samples of test questions or practice tests, directions, answer sheets, manuals, and score reports.	A-6. Evaluate representative samples of test questions or practice tests, directions, answer sheets, manuals, and score reports before selecting a test.
A-7. Avoid potentially offensive content or language when developing test questions and related materials.	A-7. Evaluate procedures and materials used by test developers, as well as the resulting test, to ensure that potentially offensive content or language is avoided.
A-8. Make appropriately modified forms of tests or administration procedures available for test takers with disabilities who need special accommodations.	A-8. Select tests with appropriately modified forms or administration procedures for test takers with disabilities who need special accommodations.
A-9. Obtain and provide evidence on the performance of test takers of diverse subgroups, making significant efforts to obtain sample sizes that are adequate for subgroup analyses. Evaluate the evidence to ensure that differences in performance are related to the skills being assessed.	A-9. Evaluate the available evidence on the performance of test takers of diverse subgroups. Determine to the extent feasible which performance differences may have been caused by factors unrelated to the skills being assessed.

B. Administering and Scoring Tests

TEST DEVELOPERS	TEST USERS
Test developers should explain how to administer and score tests correctly and fairly.	Test users should administer and score tests correctly and fairly.
B-1. Provide clear descriptions of detailed procedures for administering tests in a standardized manner.	B-1. Follow established procedures for administering tests in a standardized manner.
B-2. Provide guidelines on reasonable procedures for assessing persons with disabilities who need special accommodations or those with diverse linguistic backgrounds.	B-2. Provide and document appropriate procedures for test takers with disabilities who need special accommodations or those with diverse linguistic backgrounds. Some accommodations may be required by law or regulation.
B-3. Provide information to test takers or test users on test question formats and procedures for answering test questions, including information on the use of any needed materials and equipment.	B-3. Provide test takers with an opportunity to become familiar with test question formats and any materials or equipment that may be used during testing.
B-4. Establish and implement procedures to ensure the security of testing materials during all phases of test development, administration, scoring, and reporting.	B-4. Protect the security of test materials, including respecting copyrights and eliminating opportunities for test takers to obtain scores by fraudulent means.
B-5. Provide procedures, materials and guidelines for scoring the tests, and for monitoring the accuracy of the scoring process. If scoring the test is the responsibility of the test developer, provide adequate training for scorers.	B-5. If test scoring is the responsibility of the test user, provide adequate training to scorers and ensure and monitor the accuracy of the scoring process.
B-6. Correct errors that affect the interpretation of the scores and communicate the corrected results promptly.	B-6. Correct errors that affect the interpretation of the scores and communicate the corrected results promptly.
B-7. Develop and implement procedures for ensuring the confidentiality of scores.	B-7. Develop and implement procedures for ensuring the confidentiality of scores.

C. Reporting and Interpreting Test Results

TEST DEVELOPERS	TEST USERS
Test developers should report test results accurately and provide information to help test users interpret test results correctly.	Test users should report and interpret test results accurately and clearly.
C-1. Provide information to support recommended interpretations of the results, including the nature of the content, norms or comparison groups, and other technical evidence. Advise test users of the benefits and limitations of test results and their interpretation. Warn against assigning greater precision than is warranted.	C-1. Interpret the meaning of the test results, taking into account the nature of the content, norms or comparison groups, other technical evidence, and benefits and limitations of test results.
C-2. Provide guidance regarding the interpretations of results for tests administered with modifications. Inform test users of potential problems in interpreting test results when tests or test administration procedures are modified.	C-2. Interpret test results from modified test or test administration procedures in view of the impact those modifications may have had on test results.
C-3. Specify appropriate uses of test results and warn test users of potential misuses.	C-3. Avoid using tests for purposes other than those recommended by the test developer unless there is evidence to support the intended use or interpretation.
C-4. When test developers set standards, provide the rationale, procedures, and evidence for setting performance standards or passing scores. Avoid using stigmatizing labels.	C-4. Review the procedures for setting performance standards or passing scores. Avoid using stigmatizing labels.
C-5. Encourage test users to base decisions about test takers on multiple sources of appropriate information, not on a single test score.	C-5. Avoid using a single test score as the sole determinant of decisions about test takers. Interpret test scores in conjunction with other information about individuals.
C-6. Provide information to enable test users to accurately interpret and report test results for groups of test takers, including information about who were and who were not included in the different groups being compared, and information about factors that might influence the interpretation of results.	C-6. State the intended interpretation and use of test results for groups of test takers. Avoid grouping test results for purposes not specifically recommended by the test developer unless evidence is obtained to support the intended use. Report procedures that were followed in determining who were and who were not included in the groups being compared and describe factors that might influence the interpretation of results.
C-7. Provide test results in a timely fashion and in a manner that is understood by the test taker.	C-7. Communicate test results in a timely fashion and in a manner that is understood by the test taker.
C-8. Provide guidance to test users about how to monitor the extent to which the test is fulfilling its intended purposes.	C-8. Develop and implement procedures for monitoring test use, including consistency with the intended purposes of the test.

D. Informing Test Takers

Under some circumstances, test developers have direct communication with the test takers and/or control of the tests, testing process, and test results. In other circumstances the test users have these responsibilities.

Test developers or test users should inform test takers about the nature of the test, test taker rights and responsibilities, the appropriate use of scores, and procedures for resolving challenges to scores.

D-1. Inform test takers in advance of the test administration about the coverage of the test, the types of question formats, the directions, and appropriate test-taking strategies. Make such information available to all test takers.

D-2. When a test is optional, provide test takers or their parents/guardians with information to help them judge whether a test should be taken—including indications of any consequences that may result from not taking the test (e.g., not being eligible to compete for a particular scholarship)—and whether there is an available alternative to the test.

D-3. Provide test takers or their parents/guardians with information about rights test takers may have to obtain copies of tests and completed answer sheets, to retake tests, to have tests rescored, or to have scores declared invalid.

D-4. Provide test takers or their parents/guardians with information about responsibilities test takers have, such as being aware of the intended purpose and uses of the test, performing at capacity, following directions, and not disclosing test items or interfering with other test takers.

D-5. Inform test takers or their parents/guardians how long scores will be kept on file and indicate to whom, under what circumstances, and in what manner test scores and related information will or will not be released. Protect test scores from unauthorized release and access.

D-6. Describe procedures for investigating and resolving circumstances that might result in canceling or withholding scores, such as failure to adhere to specified testing procedures.

D-7. Describe procedures that test takers, parents/guardians, and other interested parties may use to obtain more information about the test, register complaints, and have problems resolved.

Note: The membership of the Working Group that developed the *Code of Fair Testing Practices in Education* and of the Joint Committee on Testing Practices that guided the Working Group is as follows:

Peter Behuniak, PhD
Lloyd Bond, PhD
Gwyneth M. Boodoo, PhD
Wayne Camara, PhD
Ray Fenton, PhD
John J. Fremer, PhD (Co-Chair)
Sharon M. Goldsmith, PhD
Bert F. Green, PhD
William G. Harris, PhD
Janet E. Helms, PhD

Stephanie H. McConaughy, PhD
Julie P. Noble, PhD
Wayne M. Patience, PhD
Carole L. Perlman, PhD
Douglas K. Smith, PhD (deceased)
Janet E. Wall, EdD (Co-Chair)
Pat Nellor Wickwire, PhD
Mary Yakimowski, PhD

Lara Frumkin, PhD, of the APA
served as staff liaison.

The Joint Committee intends that the Code be consistent with and supportive of existing codes of conduct and standards of other professional groups who use tests in educational contexts. Of particular note are the Responsibilities of Users of Standardized Tests (Association for Assessment in Counseling, 1989), APA Test User Qualifications (2000), ASHA Code of Ethics (2001), Ethical Principles of Psychologists and Code of Conduct (1992), NASP Professional Conduct Manual (2000), NCME Code of Professional Responsibility (1995), and Rights and Responsibilities of Test Takers: Guidelines and Expectations (Joint Committee on Testing Practices, 2000).