Bias and Fairness in State Testing

2013

Authorizing legislation: EHB 1450 (2013)

Assessment and Student Information
Robin G. Munson, Ph.D., Assistant Superintendent

Prepared by:

- Robin Munson, Assistant Superintendent (robin.munson@k12.wa.us, 360-725-6336)
- Cinda Parton, Director of Assessment Development (cinda.parton@k12.wa.us, 360-725-6291)
# TABLE OF CONTENTS

Executive Summary ......................................................................................................................... 3  
Introduction .................................................................................................................................. 4  
Bias and Fairness Activities Conducted by OSPI for Current State Tests ........................................ 4  
Bias and Fairness Activities Conducted by Smarter Balanced for Future State Tests ....................... 6  
Bias and Fairness Activities Conducted by Other States .................................................................. 8  
Conclusion .................................................................................................................................... 8  
References ..................................................................................................................................... 9  

**Appendices**

Appendix A: Bias and Sensitivity Review Guidelines ...................................................................... 10  
Appendix B: Bias and Sensitivity Training PowerPoint .................................................................... 12  

**List of Figures**

Figure 1: Test Development Process ............................................................................................ 4  

Executive Summary

Statewide testing ensures all public school students, no matter where they go to school, receive a quality education. Washington students are tested regularly as they move through school. In high school, they must pass specific exit exams, or state-approved alternatives, to be eligible to graduate.

Test results should not be influenced by a student’s disability, ethnicity, gender, or English language ability. The Office of Superintendent of Public Instruction (OSPI) and the state’s testing vendors are committed to making sure these tests are unbiased and fair. From item writing through pilot testing and psychometric studies, our tests are built to be valid and accessible by the widest possible range of students.

As our state transitions to the tests designed by the Smarter Balanced Assessment Consortium, the process will change. But the commitment to unbiased and fair testing will remain a priority.
**Introduction**

It is critically important to all students in Washington that state assessments are free of bias and are fair to students of various backgrounds. The Office of Superintendent of Public Instruction (OSPI) has a test development process that strives to prevent bias and ensure fairness. The processes used with our current state tests, as well as the steps already taken by (and future steps planned by) the Smarter Balanced Assessment Consortium, are described below.

**Bias and Fairness Activities Conducted by OSPI for Current State Tests**

OSPI's test development process includes several activities designed to prevent bias and ensure fairness in state tests. All assessments are developed in a standardized way, as illustrated in Figure 1. Bias and fairness are addressed in item writing and review activities, bias and sensitivity reviews, pilot testing, and psychometric analyses in data review.

**Figure 1. Test Development Process**

Item writing for Washington's current tests is a combined effort of our development contractor (Educational Testing Service), OSPI assessment content staff, and teachers recruited from around the state. Teacher item writers are chosen from the pool of applicants for the item writing committee so as to represent the diversity of students in Washington. Item writers are trained to write items that are free of bias and peer reviews are conducted during the item writing process to allow the entire committee to screen items for potential bias issues. Item writing committees are provided training that includes the following guidelines:

- The use of language should be straightforward, grade-appropriate, and free of idiomatic expressions.
- The inclusion of diagrams and art should be well-labeled, accurate, and selected for ease of visual discrimination.
• The choice of contexts should be familiar to examinees regardless of socioeconomic status, region, language proficiency, cultural background, or gender.

Bias and sensitivity reviews routinely occur before new items are piloted. Bias and Sensitivity Committee members are recruited from throughout the state representing various racial groups, both genders, and different sexual orientations. Members of the committee are invited to participate based upon their individual background to ensure representation of the diversity of Washington state’s population. Item writers are not involved in the committee or this review.

Committee members receive training on what to look for when reviewing items. "The Bias and Sensitivity Review Guidelines," a document used for training, is included in Appendix A. The training presentation from a recent Bias and Sensitivity event is included in Appendix B. Committee members identify any potential elements in the language and the graphics in test items and passages/scenarios that may create an unfair disadvantage to a student due to gender, sexual orientation, cultural and ethnic background, religious belief, age, socio-economic status, or disability. Each committee member is expected to read the items and passages/scenarios during the meeting and voice an opinion concerning any elements that may result in an unfair advantage or disadvantage to a specific group of students or elements that could hinder certain students from taking the test.

The outcome from this committee is suggestions by committee members for accepting, editing/modifying, or rejecting, items. These outcomes are then provided to the OSPI content assessment staff. OSPI staff make the final decision on whether to accept the revisions/modifications suggested by the Bias and Sensitivity Committee. The majority of the edits are generally accepted. Occasionally suggested edits cannot be made due to the expectations described in the learning standards being assessed.

All new items developed for Washington’s current tests are piloted before being used for operational scoring. Pilot items are embedded in the operational tests, but not included in the operational score. Embedded piloting allows the state to get students’ best performances and conduct psychometric analyses on newly developed items before they are used operationally.

The final step in the test development process, before an item becomes operational, is data review. One psychometric analysis conducted on piloted items in data review is Differential Item Functioning (DIF). DIF analyses look for systematic differences between groups of students who differ on an irrelevant dimension (i.e., gender and race). Other analyses include item difficulty, the distribution of responses for multiple choice item distractors, and the ability of the item to discriminate among students of varying ability levels.

Items that show any signs of differential functioning are examined closely by the grade-level, content-expert educators comprising the data review committee. The committee looks for any element in the item that could cause differential performance between groups of students. The committee may either reject the item due to an identified bias issue or accept the item when no cause of the differential performance can be discerned.
Bias and Fairness Activities Conducted by Smarter Balanced for Future State Tests

Washington will be administering new assessments in English language arts (ELA) and math beginning in 2014–15, to be in alignment with the state’s new content standards. The new tests are being developed by the Smarter Balanced multi-state consortium, of which Washington is a governing-state member. Smarter Balanced processes to prevent bias and assure fairness also includes item writing activities; bias and sensitivity reviews; pilot and field testing; and psychometric analyses.

Item Writing Activities

- Smarter Balanced conducts internal training on bias and sensitivity for each item author.
- Smarter Balanced systematically incorporates an internal bias and sensitivity checklist that each item author uses during item writing.
- Performance tasks require a student to read and understand contextual information, which often requires content-irrelevant terminology. Smarter Balanced continues to develop a classroom activity for each performance task. The classroom activity explains contextual information and reviews construct-irrelevant vocabulary. This will help ensure that the contextual information and vocabulary are accessible to all students.

Bias and Fairness Reviews

- Each item goes through Smarter Balanced internal review specifically designed to examine issues of bias and sensitivity.
- A sample of items goes through a Smarter Balanced external review for bias and sensitivity. Educators as well as staff from state assessment departments and curriculum and instruction department who specialize in issues of bias and fairness participate in the Smarter Balanced reviews. Reviewers receive training and use the bias and sensitivity checklist.
- The item writing vendor conducts an evaluation of comments to detect patterns that need to be addressed.

Role of Pilot Testing

- Cognitive labs were conducted with students who were representative of diverse socioeconomic conditions, English language proficiency, geographic locations, gender, and disabilities. The cognitive labs research aimed to examine item format, function, and features. The following are a few examples of the questions asked during cognitive labs:
  - What is the appropriate level of detail in item instruction for technology enhanced (TE) items? How does this vary by grade?
Under what conditions do students with lower reading ability use text-to-speech (TTS) to help focus on content in math? Is this affected by the quality of the voice pack?

For multi-part selected response (SR) items where students may select more than one answer choice, which wording best indicates to the student that he or she is allowed to select more than one option?

- For multi-part dichotomous choice items (e.g., yes/no), do students know that they need to answer each part?

How long does it take for students to read through complex tests, performance tasks, etc.? Is timing affected by the way students are presented the passage and items?

Contextual glossaries are item-specific glossaries that provide a definition of a word that is targeted to, and appropriate for, the context in which the word is used in the item. Are these a fair and appropriate way to support students who need language support?

Does presentation of longer passages reduce student engagement?

Can students use an equation editor tool the way it is meant to be used?

How do students interact with a tablet versus mechanical keyboard?

- Small scale trials examined the correlation between machine scored and hand scored item responses. The student responses represented diverse student populations.

- A student survey was distributed post-test session that included questions about engagement.

- A survey was administered to test administrators and students asking about their interaction with performance tasks and classroom activities associate with performance tasks.

**Role of Field Testing**

- At the end of 2013–14 field testing, calibration studies across item types, including analysis of student performance when having access to various designated supports and accommodations, will be conducted.

**DIF Analyses**

- Traditional DIF analysis studies will be conducted each year. DIF analyses will include comparisons for English language learners and students with disabilities as well as traditional groups (i.e., race and gender). Psychometricians who specialize in various weighing techniques for English language learners in a consortium context will be advising the consortium because states vary in their categorization of ELLs.
Other Item Review Activities

- Smarter Balanced worked with two multidisciplinary panels of experts, one per content area, to develop item quality review criteria. The panels included nationally respected researchers and practitioners who focus on content, disabilities, English language learners, engagement, and measurement.

- Item writers, item vendor reviewers, and Smarter Balanced reviewers use the item quality review criteria during item development.

- Smarter Balanced and their item vendor worked on creating, including reviewing, item specifications that specifically address item accessibility. These item specifications were used to create archetypes, model items that would serve as a foundation for full item development. Item specifications and archetypes were also reviewed by members of the item quality review panel.

Ongoing Monitoring of Bias and Fairness

- Smarter Balanced will have a certification process for educators and others who will contribute to item writing.

Bias and Fairness Activities Conducted by Other States

Because all state assessment programs are reviewed by the United States Department of Education (USED), through the No Child Left Behind Peer Review process, states generally follow the test development and analytic practices described in each of the two previous sections. The USED's Standards and Assessments Peer Review Guidance: Information and Examples for Meeting Requirements of the No Child Left Behind Act of 2001, includes items related to ensuring tests that are fair and free from bias. Peer reviewers look for evidence that “the State has taken appropriate judgmental (e.g., committee review) and data-based (e.g., bias studies) steps to ensure that its assessment system is fair and accessible to all students [and that] review committees have included representation of identified subgroups.”

Washington’s comprehensive assessment program received full peer review approval in 2008.

The peer review process was temporarily paused in 2012 pending transition to the new assessments of Common Core State Standards.

Conclusion

Washington’s state assessment program includes several activities to ensure fairness and prevent bias in student tests. Both our current and future tests are built to be valid and accessible for use by the widest possible range of students, including students with disabilities and students with limited English proficiency. From item writing through pilot testing and psychometric studies, OSPI and the state’s testing vendors are committed to making sure test results are not influenced by disability, ethnicity, gender, or English language ability.
References
Appendix A: Bias and Sensitivity Review Guidelines

These guidelines list a number of considerations for you to think about as you move forward in the bias and sensitivity review process. It is not the intent of the panel to discourage the teaching of controversial issues or issues which bring forth strong emotion. However, since there can be no teacher intervention during the testing process, the review committee will advise on edits or elimination of such elements in the reading passages, science scenarios, or test items. The list below is not conclusive and exists only to assist the committee in its discussion.

1. Race/Ethnic/Cultural Considerations
   - Does the material favor one racial or ethnic group over others?
   - Does the material portray one or more ethnic groups or culture in a negative or stereotypical manner or trivialize any group?
   - Does the material use language, content, or context that is not accessible or not widely familiar to one or more racial or ethnic groups or culture?
   - Does the material contain language or content that contradicts values held by a certain culture?
   - Does the material trivialize significant or tragic human experiences?

2. Sex Considerations
   - Does the material use language, content, or context that is offensive to males or females?
   - Does the material use language, content, or context that is not accessible to or not widely familiar to either males or females?
   - Does the material negatively represent or stereotype people based on gender or sexual preference?
   - Does success with the material assume that the examinee has experience with a certain type of family structure?
   - Does the material present sexual innuendoes?

3. Religious Considerations
   - Does the material favor one religion and/or demean others?
   - Does the material portray one or more religious leaders in a negative or stereotypic manner?
   - Does the material use language, content, or context that is offensive to one or more religious groups (for example, reference to witchcraft, sorcery and magic)?
   - Does the material use religious language, content, or context that is not commonly understood by members of all religious groups?
• Does the material require the parent, teacher, or examinee to support a position that is contrary to their religious beliefs?
• Does the material require knowledge of individuals, events, or groups that is not familiar to all students?

4. **Age Considerations**

• Does the material favor one age group over others except in a context where experience or maturation is relevant?
• Does the material portray one or more age groups in a negative or stereotypical manner?
• Does the material use language, content, or context that is offensive to one or more age groups?
• Does the material use language, content, or context that is not accessible to one or more of the age groups tested?

5. **Disability Considerations**

• Does the material degrade people on the basis of physical appearance or of any physical, cognitive, or emotional challenge?
• Does the material focus only on a person’s disability rather than portraying the whole person?
• Does the material use language, content, or context that is offensive to people with disabilities?
• Does the material portray one or more people with disabilities in a negative or stereotypical manner?

6. **Socio-Economic Considerations**

• Does the material suggest that affluence is related to merit or intelligence?
• Does the material use language, content, or context that is offensive to people of a particular economic status?
• Does the material favor one socioeconomic group over another?
• Is a particular ethnic, cultural, or religious group stereotyped as belonging to a specific socioeconomic status?
• Does the material romanticize or demean people based on socioeconomic status?
• Does the material suggest that poverty is related to increased negative behaviors in society?

7. **Other Considerations**

• Does the material require a student to take a position that challenges parental authority?
• Does the material present war or violence in an overly graphic manner?
• Does the material present inflammatory or highly controversial themes such as death, wars, abortions, or euthanasia, except where they are needed to meet State Content Standards?
• Does the material assume values not shared by all test takers?
• Does the material degrade people or cultures from certain regions of the country or state?
• Does the material accept or fail to denounce criminal, illegal, or dangerous behavior?
• Does the material require examinees to disclose value that they would rather hold confidential?
• Does the material contain harassing or homophobic language?

**Appendix B: Bias and Sensitivity Training PowerPoint**

OSPI provides equal access to all programs and services without discrimination based on sex, race, creed, religion, color, national origin, age, honorably discharged veteran or military status, sexual orientation including gender expression or identity, the presence of any sensory, mental, or physical disability, or the use of a trained dog guide or service animal by a person with a disability. Questions and complaints of alleged discrimination should be directed to the Equity and Civil Rights Director at (360) 725-6162 or P.O. Box 47200 Olympia, WA 98504-7200.

Download this material in PDF at http://www.k12.wa.us/LegisGov/Reports.aspx. This material is available in alternative format upon request. Contact the Resource Center at (888) 595-3276, TTY (360) 664-3631. Please refer to this document number for quicker service: 13-0049.