

# Considerations in choosing an Assessment

Assessment Purpose

Responsiveness to all Stakeholders

Inclusion in Accountability Systems

Utility of Data/Reporting System

Influence on Instructional Decisions

Assessment/Implementation Costs

Federal Approval

Assessment Quality

# Assessment Quality Framework for Utah

## 1. Validity

- a. Is there technical evidence that the assessment is valid? (reported outcome is accurate, the test determines what it is intended to determine)
- b. Are the performance levels consistent with what the state expects/requires?
- c. Do students scoring well on the assessment also perform well on assessments designed to measure the same domain?

## 2. Reliability

- a. Is there evidence that the assessment is reliable? (similar scores for similar students)
- b. Is the conditional standard error of measure sufficient?

## 3. 3<sup>rd</sup> Party Alignment

- a. Does the assessment meet industry standards using a “two-way” alignment (standards to questions and questions to standards) showing alignment to core standards?
- b. Is there evidence of appropriate depth of knowledge, range of knowledge and balance of representation of questions that reflect the Utah core?

## 4. Item Quality

- a. Is there evidence of bias and sensitivity reviews that includes consideration of the diversity of the tested population (ethnic groups, students with diverse linguistic backgrounds and students with disabilities)?
- b. Is there evidence of responsiveness to instruction, question clarity and lack of cueing?
- c. Does the assessment use multiple question types? (multiple choice, constructed response, etc.)

## 5. Fairness

- a. Is there adherence to Universal Design (UD) and Accessibility by Design principles?
- b. Does the assessment include accommodations to allow participation of all students to the fullest extent possible? (Including allowing for all accommodations that are available during instruction.)
- c. Are multiple test versions available? (translations, braille, etc.)

## 6. Technical Review/Oversight

- a. Are the technical aspects of the assessment system regularly reviewed (at least once every year) by a committee of independent measurement and related experts to provide advice on how best to improve the quality of the assessments?

## 7. Administration

- a. Are there standardized test administration procedures?
- b. Is the number of tests and length of testing time appropriate for the desired uses?

## 8. Sustainability

- a. Is the assessment system affordable?
- b. Is there evidence that the assessment provider/contractor has consistently produced high quality products, is responsive to customer needs, and can meet budgets and deadlines?

## **Three Options for Future Assessment Systems**

1. Procurement of an existing computer adaptive assessment
2. Procurement for state developed computer adaptive assessment
3. Multi-state consortium developed computer adaptive assessment

## **Procurement of an Existing Computer Adaptive Assessment**

### **Description**

An existing computer adaptive assessment is an assessment that is developed and produced by an assessment vendor. There are an increasing number of vendors who are developing computer adaptive tests. Examples would include NWEA, Pearson, SRI and ALEKS. The assessment would include the adaptive test, the testing engine and the reporting system.

### **Procedure**

- USOE, in collaboration with stakeholders, would develop a list of requirements for the assessment.
- An RFP would be created and posted
- A committee would review respondent's proposals
- State purchasing would make the final vendor selection decision
- USOE would complete contract negotiations and definition of scope of work

### **Considerations**

- All features of the assessment would be known at the beginning of the contract
- The assessment system and reporting system may not be customizable to adjust to future state needs
- Current systems are limited to only multiple choice items
- Limited control of vendor selection, price driven decision process  
Examples: WIDA, Online Writing, Idaho SAT
- Limited ability to respond to 3<sup>rd</sup> party alignment results
- Costs would be fixed and could increase
- Procurement constraints require a re-issuance of an RFP at least every five years, that is, the test could change upon re-issuance  
Example of impact: Iowa/SAT NRTs
- Would require a federal peer review process
- May or may not be comparable to students in other states

# **Procurement for a State Developed Computer Adaptive Assessment**

## **Description**

The procurement for a state developed computer adaptive assessment would require USOE to contract with an assessment vendor to assist the USOE in developing assessment questions, the adaptive algorithms, the testing engine, and the reporting system that is specifically designed around Utah specifications. This assessment would include the adaptive test, the testing engine and the reporting system.

## **Procedure**

- USOE, in collaboration with stakeholders, would develop a list of requirements for the assessment
- USOE would analyze existing test questions, technical resources, etc., to determine RFP requirements
- An RFP would be created and posted
- A committee would review respondent's proposals
- State purchasing would make the final vendor selection decision
- USOE would complete contract negotiations and definition of scope of work

## **Considerations**

- Due to the development process, the completed assessment would not be visible at the beginning of the contract
- The assessment and reporting system would be customizable to adjust to future state needs
- Test questions could include multiple choice, constructed response and technology enhanced
- Limited control of vendor selection, price driven decision process
- Collaborative design process to meet Utah's needs would include Utah educators
- 3<sup>rd</sup> party alignment information could result in adjustments to the assessment
- Existing questions and resources could be leveraged
- Accommodations/accessibility would be designed to meet specific Utah requirements/needs
- Potential reduction in costs because software code and developed item bank are owned by USOE
- Would require a federal peer review process
- Is not comparable to students in other states

## **Multi-state Consortium Developed Computer Adaptive Assessment SMARTER Balanced Assessment Consortium (SBAC)**

### **Description**

SBAC will be a comprehensive assessment system that includes formative tools (web based tools for teachers); interim tests (optional computer adaptive tests that teachers can administer whenever they teach a specific standard or unit); summative tests (computer adaptive tests and performance tasks). The computer adaptive test will include multiple choice, constructed response and technology enhanced questions. A comprehensive reporting system will provide immediate results.

### **Procedure**

- Assessment system will be implemented in the school year 2014/15
- State MOUs are not binding

### **Considerations**

- The assessment is in the development process, currently there is not a test to review
- Utah is participating in test question development (multiple choice, constructed response, technology enhanced items, performance tasks)
- The system will include many adaptive tests available for local use
- Higher education partnership focused on not requiring remediation for students
- The multiple state administration will allow student scores to be compared to students in other states
- Growth scores will be calculated for students who move from one state to another.
- The assessment will be free to all states, but there will be costs for implementation
- Will not require a federal peer review process