The CAAVES Project

- Consortium involves 6 states: AZ, HI, ID, IN, NV, and MS.

- A Project Goal included an experimental study to investigate feasibility of item modification strategies for future alternate assessments.

- Experimental study designed with a cognitive lab and a computer based delivery system of a common set of original and modified items.
The project’s goal will be accomplished by...

1. Developing a common set of test items from existing reading and mathematics tests using modification principles that facilitate reading access and valid responses and

2. Using a computer-based delivery system to experimentally examine student preferences, score comparability, and item statistics of the modified items for students with and without disabilities

3 Groups of Students Selected across 5 states

1. General Education Students
2. Students with Disabilities who do not meet the participation criteria set forth in the federal regulations
3. Students with Disabilities who meet the participation criteria outlined in the federal regulations
3 Criteria for the Students with Disabilities MAA Eligible group

200.1(e)(2) In the guidelines that a State establishes under paragraph (f)(1) of this section, criteria must include, but are not limited to, each of the following:

The student:
1. Has an IEP with goals based on academic content standards for the grade enrolled.

The student...
2. Has a disability precluding the student from achieving grade-level proficiency, as demonstrated by the student’s performance on the state assessment or another assessment that documents academic achievement.

3. Progress to date (a) in response to appropriate instruction, is addressing the student’s individual needs and (b) based on multiple measurements is such that, even if significant growth occurs, the student will not achieve grade-level proficiency within the year covered by the student's IEP.
**Item Modification Procedures**

- **Goals of the Consortium Workgroups:**
  1. To modify existing items in reading and mathematics so they are less complex, more accessible, and likely easier yet still measure the same knowledge and skill as the original items.
  2. Document item modification procedures and guiding principles, challenges, and recommendations for future work.

**Started with a Menu**

- Remove a response option
- Simplify language & improve readability
- Add graphic support
- Reorganize the layout
Test Items from 4 Content Strands

- Reading Comprehension
- Vocabulary
- Numbers & Operations
- Data Analysis & Probability

Before & After Comprehension & Vocabulary

Read the following passage and then answer questions 1 and 2.

As I tightened my seat belt and prepared for take off, I could feel my heart pounding. But even if it was the first time flying, I was more calm than Jeff. After we took off, I was able to relax. After the short flight, I asked Jeff if he felt it would have been as smooth as the flight we had taken before.

1. Based on this passage, we could predict that the sentence
   "I will have a fear of flying"
   is
   A. will have less fear of flying
   B. will never fly again
   C. will always fly with hands
   D. will never fly without a parachute

2. Jeff was already ready for the flight before. He had given the plane a final check.
   "Place of balance" is another
   A. will have less fear of flying
   B. will never fly again
   C. will always fly with hands
   D. will never fly without a parachute

3. Jeff was already ready for the flight before. He had given the plane a final check.
   "Place of balance" is another
   A. will have less fear of flying
   B. will never fly again
   C. will always fly with hands
   D. will never fly without a parachute

4. Jeff was already ready for the flight before. He had given the plane a final check.
   "Place of balance" is another
   A. will have less fear of flying
   B. will never fly again
   C. will always fly with hands
   D. will never fly without a parachute

5. Jeff was already ready for the flight before. He had given the plane a final check.
   "Place of balance" is another
   A. will have less fear of flying
   B. will never fly again
   C. will always fly with hands
   D. will never fly without a parachute

6. Jeff was already ready for the flight before. He had given the plane a final check.
   "Place of balance" is another
   A. will have less fear of flying
   B. will never fly again
   C. will always fly with hands
   D. will never fly without a parachute

7. Jeff was already ready for the flight before. He had given the plane a final check.
   "Place of balance" is another
   A. will have less fear of flying
   B. will never fly again
   C. will always fly with hands
   D. will never fly without a parachute

8. Jeff was already ready for the flight before. He had given the plane a final check.
   "Place of balance" is another
   A. will have less fear of flying
   B. will never fly again
   C. will always fly with hands
   D. will never fly without a parachute

9. Jeff was already ready for the flight before. He had given the plane a final check.
   "Place of balance" is another
   A. will have less fear of flying
   B. will never fly again
   C. will always fly with hands
   D. will never fly without a parachute

10. Jeff was already ready for the flight before. He had given the plane a final check.
    "Place of balance" is another
    A. will have less fear of flying
    B. will never fly again
    C. will always fly with hands
    D. will never fly without a parachute
Before & After – Numbers

4. Mr. White is buying a new car and wants to make a down payment of $1,750.00. When he withdraws this amount from his savings account, the teller tells him that the largest bills that he can get are $100 bills. Mr. White asks for as many hundred dollar bills as possible. How many hundred dollar bills should he receive?

- A. 17
- B. 16
- C. 7
- D. 6

Before & After – Data Analysis

3. A car dealer sold 360 cars last month. The graph shows the percentage of white, black, red, and yellow cars sold.

- About how many were black?
  - A. 190
  - B. 130
  - C. 85
  - D. 20

3. The graph shows the percentage of cars that were sold by color.

- Color of Cars Sold
  - 11% yellow
  - 10% red
  - 32% black
  - 47% white

- About how many black cars were sold?
  - A. 190
  - B. 130
  - C. 85
  - D. 20
8. Michael’s quiz scores are 50, 86, 66, 60, and 40. What must he make on the next quiz in order to have a 70 quiz average?

- A. 14
- B. 84
- C. 90
- D. 96

8. Mike’s test scores are 50, 86, 66, 60, and 40. What score does he need on the next test in order to have a mean of 70?

- A. 96
- B. 90
- C. 84

---

**Before & After – Analyze Item Statistics**

- Designed by Beddow, Kettler, & Elliott for the CAAVES project.
- Purpose: to provide an organized framework for item modification.
- Closely follows guidelines from NCEO Technical Report #42. (Thompson, Johnstone, Anderson, & Miller, 2006)
Theoretical Influences

- **Universal Design**: The design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design. (Center for Universal Design, NC State University, 2007)

- **Cognitive Load Theory**: Learning principles that result in efficient instruction by leveraging cognitive processes. (Clark, Nguyen, & Sweller, 2006)

Universal Design

- Equitable use
- Flexibility in use
- Simple and intuitive
- Perceptible information
- Tolerance for error
- Low physical effort
- Size and space for approach and use

(Center for Universal Design, 2007)
Cognitive Load Theory

- Intrinsic Load
- Extraneous Load
- Germane Load

Learning

Organization

- Two sections:
  - Considerations for All Universally Designed Assessment Items
    (7 categories, 38 considerations)
  - Considerations for Computer-Based Tests
    (5 categories, 26 considerations)

- First, the rater provides a brief description of the construct the item is intended to measure.

- Then, for each accessibility consideration, the rater is asked to check Yes, No, or N/A.
Finally, the rater is asked to circle the number that best represents his/her summative evaluation of the item’s accessibility for that category based on checklist responses:

<table>
<thead>
<tr>
<th>Number</th>
<th>Accessibility Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>The item is Not Accessible</td>
</tr>
<tr>
<td>1</td>
<td>The item is Minimally Accessible</td>
</tr>
<tr>
<td>2</td>
<td>The item is Moderately Accessible</td>
</tr>
<tr>
<td>3</td>
<td>The item is Maximally Accessible</td>
</tr>
</tbody>
</table>

### Categories of the IAMG

<table>
<thead>
<tr>
<th>All Tests</th>
<th>Computer-Based Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content Relevance</td>
<td>Layout and Design</td>
</tr>
<tr>
<td>Text Content</td>
<td>Navigation</td>
</tr>
<tr>
<td>Text Economy/Conciseness</td>
<td>Screen Reader Considerations</td>
</tr>
<tr>
<td>Text Appearance</td>
<td>Test Specific Options</td>
</tr>
<tr>
<td>Visual Content</td>
<td>Computer Capabilities</td>
</tr>
<tr>
<td>Fairness for Subgroups</td>
<td></td>
</tr>
<tr>
<td>Format Flexibility</td>
<td></td>
</tr>
</tbody>
</table>
Item Modification
Cognitive Lab Overview

Our study involved 3 components:

- Students completed a series of 16 assessment items (8 reading; 8 math).
- Students were asked to think aloud as they completed or solved these items.
- We also asked follow-up questions about students’ perceptions of the assessment items.

(Johnstone, Bottsford-Miller, & Thompson, 2006; Branch, 2000).

Test Item Modification Distribution

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test A</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

X = Item modifications used.
### Sample Size by Sub-group

<table>
<thead>
<tr>
<th></th>
<th>Test A group</th>
<th>Test B group</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students without disabilities</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Students with disabilities; Testing Accommodations</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Students with disabilities; Modified Alternate Assessment</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

### Procedures Overview

- We explained the think-aloud procedures, had the students restate their understanding of the process, and modeled thinking aloud on a practice item.
- We used a script adapted from a study conducted by Johnstone, Bottsford-Miller, and Thompson, 2006.
- Students were prompted only when they were silent for 10 consecutive seconds.
- If students verbalized infrequently, we reminded them to “keep thinking aloud” or “keep talking.” Otherwise we generally did not give encouragement or support.
Results: Visuals and Graphs

- Most SWDs (67%) saw the visuals as being helpful and providing support on reading questions and passages.
- 100% of the students without disabilities indicated the pictures made no difference in understanding the reading questions or passages.
- Students with (50%) and without disabilities (67%) generally saw the visuals and graphs as being helpful and providing support on math items.
- ...However, 33% of SWDs indicated that the visuals/graphs were distracting or made it harder to answer the questions.

Results: Directions and Bold Type

- General education students (67%) and student eligible for MAAs (100%) generally preferred test directions that were explicit, such as “Read...then answer....”
- Some students indicated that the less explicit directions (i.e., “Use the passage...”) might encourage test takers to skim rather than read closely.
- The majority of students from all groups (78% of the total) felt the use of bold type to identify key terms was helpful in answering the reading items.
Results: 3 vs. 4 Answer Choices

- SWDs (with one exception) perceived no difference in difficulty between items having 3 or 4 possible answers on reading items.
- Conversely, 67% of the students without disabilities identified the 3-answer modification as making the reading items easier.
- The results suggest that this modification did not affect either groups' performance on most reading items [e.g., only one item ("Pesticides") had a discernable difference in student accuracy between modified and unmodified versions].

Unmodified Item

Pesticides

In the late 1980s, farmers began to use a pesticide to control insects that harmed their cotton crops. This problem was solved. However, an insect group that pollinated the corn crops was also injured. Without pollination the corn kernels did not fully develop. This affected the corn harvest on which the farm families had come to depend. What is not mentioned as one effect of pesticide usage?

A. soil contamination
B. destruction of pests
C. destruction of friendly insects
D. crop losses
Modified Item

Pesticides
In the late 1980s, farmers began to use a chemical pesticide. It was used to control insects that harmed their cotton crops. This solved one problem, but caused another. An insect group that pollinated the corn crops was also harmed by the pesticide. Without pollination the corn kernels did not fully develop. This decreased the corn harvest.

What is NOT mentioned as one effect of using chemical pesticides?

A. destruction of the soil
B. destruction of pests
C. destruction of friendly insects

Results:
3 vs. 4 Answer Choices

- Students without disabilities (67%) and SWDs in TA group (67%) generally indicated 3 answer choices made the math items easier.
- Some students in these groups appeared to use the possible answer choices to help solve math items, but it was not clear that they used this same strategy in reading.
- For the MAA group, the 3-answer choice modification was less likely to be identified as helpful, but it did seem to make a difference on one particular item (i.e., “scientific notation”).
Results: Analogies

- Most students (including 2/3 of SWDs) found the traditional format for the analogy easier (i.e., "meteor:space::dolphin:______"). Some students indicated they had been taught analogies using this format and it was familiar to them.

- This was supported by the results as SWDs correctly answered all the traditional analogy items. SWDs missed items with a modified analogy format (i.e., meteor is to space as dolphin is to ___) 40% of the time.

- Familiar test item format may help make an item less difficult.

Take Away Ideas

- "Conservative" modifications were used and the effects (on student performance) generally were modest. More “aggressive” modifications might result in more robust effects.

- SWDs sometimes appeared unfamiliar with concepts (e.g., percentages) or incorrectly applied problem solving strategies. In these cases, item modifications are unlikely to provide support or facilitate access.

- Reading fluency may be an issue for SWDs. Some SWDs (in the MAA group) took up to 6 minutes to read short reading passages, resulting in testing sessions that were almost twice as long as their peers. How could (or should) technology be used to address this barrier?
References


Experimental Study Progress

- November 2007- Consortium members reviewed Cognitive Lab Results
- Lessons Learned to date:
  - Modifications appear too conservative
  - Reflect & consider Cognitive Load Theory
  - Participation criteria & identification process takes time
  - Involve general education content specialists, special education specialists, and measurement/assessment specialists in the item modification process
Experimental Study Progress

- New test design: 13 original, 13 modified without read aloud, 13 modified with read aloud
- Random assignment of students to a sequence of the items
- Read aloud feature will have a voice read directions and question stems. However, the reading items may or may not use this feature for the question stem.
- Assessment window February 2008
- Preliminary results – June 2008 at National Assessment Conference in Orlando

Any Questions?

- Elizabeth Compton, Idaho Department of Education & Boise State University, ebermen@boisestate.edu
- Ryan J. Kettler, Vanderbilt University, ryan.j.kettler@vanderbilt.edu
- Andrew T. Roach, Georgia State University, aroach@gsu.edu
- CAAVES Website: www.vanderbilt.edu/caaves/html