Text Readers for Everyone on All Tests – Getting a Handle on What This Means

June 27, 2017 • Austin, Texas
Hosted by:

- ASES SCASS
- EL SCASS
- NCEO
Do you ever wonder:

- What is the evidence of the validity of results when text readers are used by different students?
- Can we develop policies and a common language for text readers and text to speech for all students?
- Should language translations be provided through text to speech options?
...then this forum is for you!

Our goal this afternoon is to begin to develop some clarity on the implementation of text readers for ALL students, and to discuss the issues surrounding the use of text readers for ALL students.
Agenda

3:30 Introduction
   Sandra Warren, CCSSO

3:45 Overview of the Issues
   Martha Thurlow, NCEO

4:00 Results of State Policy Analysis
   Sheryl Lazarus and Kathy Strunk, NCEO

4:15 The Issues from the Perspective of:
   a researcher: Cara Laitusis, ETS
   a vendor: Peter Ramsdell, TextHelp
   a State: Melissa Gholson, West Virginia Department of Education

5:00 Break

5:10 Facilitated Breakout Discussions
   General education students
   English learners
   Students with IEPs and 504 plans

6:00 Debrief
In Your Folders

- Pre-session agenda
- Discussion questions
- Evaluation form
Overview
Goals of Session

- Identify research needs for text to speech.
- Begin to develop common terminology and definitions that more precisely describe the types of text to speech so that these descriptions can be used in RFPs.
- Discuss implementation and policy challenges and ways to address them.
Some Questions We Will Discuss

- What is the evidence of the validity of results when used by different students?
- Can we develop policies and a common language for text readers and text to speech for all students to counteract the current situation where these terms mean different things to different people (e.g., recorded human voice, synthesized voice, text highlighted or not)?
Questions – Cont.

- Should language translations be provided through text to speech options?
- What resources and training do educators, parents, and students need to support better decision making and implementation of text readers and text to speech?
Desired Outcome

- Produce a report that summarizes the session and the discussion.
  - Notetakers’ notes will be transcribed and summarized.
  - All comments will be anonymous.
  - Report will be reviewed by facilitators and notetakers before publication.

- Draft RFP language
State Policy Analysis

Sheryl Lazarus, NCEO
Kathy Strunk, NCEO
Purpose

To provide a snapshot of how accessibility features and accommodations that used technology to deliver an assessment orally are included in state policies.

- Analyzed reading/English language arts (ELA) and math content assessments
- 2016-17 school year
Tiers of Accessibility

Universal Features
for all students

Designated Features
for students who need them as identified by an educator in advance

Accommodations
for students with disabilities; in some cases, ELs are also eligible
### Number of State Policies

#### Text to Speech: *Directions*

<table>
<thead>
<tr>
<th>Accessibility Tier</th>
<th>Reading/ELA</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Student w. Disability</td>
<td>EL</td>
</tr>
<tr>
<td>Designated Feature</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>Accommodation</td>
<td>18</td>
<td>16</td>
</tr>
</tbody>
</table>

Preliminary Findings: Report Forthcoming
### Number of State Policies

**Text to Speech: Items**

<table>
<thead>
<tr>
<th>Accessibility Tier</th>
<th>Reading/ELA</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Student w. Disability</td>
<td>EL</td>
</tr>
<tr>
<td>Designated Feature</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>Accommodation</td>
<td>17</td>
<td>14</td>
</tr>
</tbody>
</table>

Preliminary Findings: Report Forthcoming
## Number of State Policies

### Text to Speech: **Passages**

<table>
<thead>
<tr>
<th>Accessibility Tier</th>
<th>Reading/ELA</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Student w. Disability</td>
<td>EL</td>
</tr>
<tr>
<td>Designated Feature</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Accommodation</td>
<td>36</td>
<td>13</td>
</tr>
</tbody>
</table>

Preliminary Findings: Report Forthcoming
Wide Variation in How Text to Speech is Included in Policies

- For example:
  - Features and policies may differ within a consortia depending on vendor used by state
  - Some policies indicate that the student can control the speed and volume of text-to-speech – others that it must be pre-determined
  - Text to speech may be provided in a different way on a Spanish translation of an assessment than how it is provided on the English version.

- Large differences in the “grain size” regarding the information provided in policies. Some provided extensive detail about when and how it should be used.; others provided little detail.
Presentations from Three Perspectives:

• A Researcher: Cara Laitusis (ETS)
• A Vendor: Peter Ramsdell (TextHelp)
• A State: Melissa Gholson (West Virginia Department of Education)
Perspectives from Research

Cara Laitusis
Overview

- What do we know?
  - Overview of recent meta-analyses
  - Summary of findings
- What are the challenges of conducting research?
- Future research studies
- What would you do?
What do we know?

Numerous studies examining “read aloud”

- 3 meta-analysis of “differential boost” in last 3 years

- Numerous studies comparing psychometric comparability

- Few studies examining predictive validity
Summary of Buzick & Stone Meta-Analysis (2014)

- 22 studies on reading
- 26 studies on mathematics
- Among the studies that were suitable to include in the meta-analysis:
  - 9 comparisons in reading
  - 20 comparisons in math
Variety of Audio Presentation

- Delivery options
  - Text to speech
  - Prerecorded text to speech
  - Prerecorded human audio
  - Human proctor (scripted or on the fly)
  - Individual vs. group administration
  - Student-paced vs. computer/proctor-paced
  - Ability to reread (i.e., words, phrases, sentences)
Variety of Audio Presentation

- Amount of text read aloud
  - Passages
  - Parts of passages/questions
    - proper nouns on ELA test
    - no numerals or symbols on math test
  - Questions and answer options
  - Directions
Reading
Summary of comparisons

<table>
<thead>
<tr>
<th>Grade</th>
<th>Elementary school (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Middle school (4)</td>
</tr>
<tr>
<td></td>
<td>Middle &amp; High (1)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of audio presentation</th>
<th>Human reader (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Audio CD (2)</td>
</tr>
<tr>
<td></td>
<td>Text-to-speech (2)</td>
</tr>
<tr>
<td></td>
<td>Video (1)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Extra time</th>
<th>Accommodated only (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standard and accommodated (3)</td>
</tr>
<tr>
<td></td>
<td>None/unknown (3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item types</th>
<th>Similar to state assessment (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Other (4)</td>
</tr>
</tbody>
</table>

8 of 22 studies included in Smarter Balanced Literature Review
## Math

### Summary of comparisons

<table>
<thead>
<tr>
<th>Grade</th>
<th>Elementary school (10)</th>
<th>Middle school (8)</th>
<th>High school (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of audio</strong></td>
<td>Human reader (8)</td>
<td>Video (10)</td>
<td>Text-to-speech (2)</td>
</tr>
<tr>
<td><strong>presentation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Extra time</strong></td>
<td>Accommodated only (2)</td>
<td>Standard and</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>accommodated (13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>None/unknown (5)</td>
<td></td>
</tr>
<tr>
<td><strong>Item types</strong></td>
<td>Similar to state</td>
<td>CBM or performance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>assessment (17)</td>
<td>assessment (3)</td>
<td></td>
</tr>
</tbody>
</table>

14 of 35 studies included in Smarter Balanced Literature Review
<table>
<thead>
<tr>
<th>Study</th>
<th>Grade</th>
<th>Sample size</th>
<th>Weight</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crawford &amp; Tindal (2004)</td>
<td>4-5</td>
<td>76</td>
<td>6.0%</td>
<td>V</td>
</tr>
<tr>
<td>Fletcher (2006)</td>
<td>3</td>
<td>47</td>
<td>2.0%</td>
<td>H</td>
</tr>
<tr>
<td>Laitusis (2010)</td>
<td>4</td>
<td>527</td>
<td>40%</td>
<td>CD</td>
</tr>
<tr>
<td>Olson &amp; Dirir [KY] (2010)</td>
<td>4</td>
<td>150</td>
<td>10%</td>
<td>H</td>
</tr>
<tr>
<td>Fletcher (2009)</td>
<td>7</td>
<td>56</td>
<td>2.0%</td>
<td>H-r</td>
</tr>
<tr>
<td>Laitusis (2010)</td>
<td>8</td>
<td>376</td>
<td>34%</td>
<td>CD</td>
</tr>
<tr>
<td>Meloy et al. (2002)</td>
<td>6-8</td>
<td>29</td>
<td>1.0%</td>
<td>H</td>
</tr>
<tr>
<td>Thurlow (2012)</td>
<td>6 &amp; 8</td>
<td>44</td>
<td>3.0%</td>
<td>Pen</td>
</tr>
<tr>
<td>Higgins (2005)</td>
<td>age 10-18</td>
<td>30</td>
<td>3.0%</td>
<td>Pen</td>
</tr>
<tr>
<td>Summary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TTS Meta-analysis (2017)

Bielinski et al. (2001) 0.28 [0.16, 0.39]
Boyle et al. (2003) 0.65 [0.01, 1.30]
Fasting & Lyster (2005) 0.90 [0.35, 1.48]
Fletcher et al. (2006) 0.92 [0.47, 1.33]
Fletcher et al. (2009) 0.50 [0.00, 1.00]
Lange et al. (2006) 0.77 [0.20, 1.33]
Lundberg, I., & Olofsson, Å. (1993) -0.32 [-1.28, 0.64]
Meloy et al. (2002) 1.10 [0.57, 1.63]
Roberts et al. (2013) 0.45 [0.06, 0.84]
Shany & Biemiller (1995) 0.82 [-0.08, 1.71]
Dolan et al. (2005) 1.02 [-0.47, 2.50]
Elkind et al. (1996) 0.10 [-0.63, 0.83]
Floyd & Judge (2012) 1.09 [0.46, 1.72]
Higgins & Raskind (1997) -0.12 [-0.39, 0.15]
Hodapp et al. (2007) -0.74 [-1.10, -0.38]
Kosciolek & Ysseldyke (2000) 0.53 [0.00, 1.06]
Laitusis (2010a) 0.57 [0.56, 0.58]
Laitusis (2010b) 0.32 [0.32, 0.32]
McCove (2012) 0.53 [-0.61, 1.67]
Meyer & Bouck (2014) -0.21 [-0.95, 0.53]
Schmitt et al. (2011) 0.23 [-0.27, 0.72]
Sorrell et al. (2007) -0.16 [-0.71, 0.39]
Thurlow et al. (2010) -0.32 [-0.53, -0.12]

RE Model 0.35 [0.14, 0.56]

Standardized Mean Difference

-2.00 -1.00 0.00 1.00 2.00 3.00
Taken together, this suggests that text-to-speech technologies may assist students with reading comprehension. (Wood, et al. 2017)

more studies are needed to further explore the moderating variables of text-to-speech and read-aloud tools’ effectiveness for improving reading Comprehension (Wood et al., 2017)

We conclude that we can expect larger score gains on average from the read aloud accommodation on reading assessments than on math assessments (Buzick & Stone, 2014)

but there is a substantial amount of variation in effect sizes that is due to factors beyond measurement error (Buzick & Stone, 2014)
What do we know?

- Read aloud improves scores, but effect sizes are larger for reading than for math for both student groups
  - Human reader had larger effect size (Li, 2014)
  - Early grades had larger effect size (Buzick & Stone, 2014)

- Differential performance gains are found on reading tests, but no/minimal differential performance gains are found in mathematics
What do we know?

- Scores with and without read aloud are psychometrically comparable* (i.e., Reliability, Factor Structure, and Differential Item Functioning)
  - *Nearly all of these studies were done with human readers or prerecorded audio (see Laitusis & Cook, 2008 for summary)
  - Studies should be replicated for CBT and TTS and can be done with operational test data

- Limited research on predictive validity but research on teachers ratings and college readiness is conflicting
  - Lessons learned from PARCC TTS Predictive Validity Study
Challenges

- Defining the construct
  - Does reading include decoding of print by touch or sight? If so,
    - How long do you have to decode?
    - Can a reading fluency test supplement comprehension tests?
  - What are we predicting?
    - Independent comprehension of text?
    - College grades?
    - Teachers ratings?
Challenges

- What grade level does comprehension overtake decoding?
- Item level accommodation use does not fit the old paradigm of “differential boost” or “interaction hypothesis”
- Few studies on predictive validity
Future Research Studies

- For the most part.... forget about differential boost for isolated accommodations/tools
- Focus on predictive validity of total score
  - Tracking students to college/career
    - Employment/High School Graduation/College Graduation rates
    - Grades in entry level courses
  - Predicting other variables
    - Future test scores
    - High School grades
    - Teacher Ratings
Future Research Studies

- Descriptive analysis of operational test data and surveys
  - Longitudinal accommodation use
  - Click stream data by item
    - Was assigned tool/feature used?
    - What other tools and features were used with TTS?
    - Do students who use AIM also use TTS tool?
    - Are there performance differences for students who use different TTS as an instructional accommodation?
What would you do?

- Change policy to only allow independent forms of read aloud on reading test after grade 6
  - Text to Speech or Screen Reader
- Include supplemental measure of oral reading fluency
  - when read aloud is provided on reading tests in grades 3-5
- Capture data on
  - Longitudinal accommodation use (for predictive validity studies)
  - Specific accommodation use by item and item elements (passage, stem, options) and make it available to researchers
  - Student survey on instructional accommodations
Questions/Comments:
Cara Laitusis
claitusis@ETS.org

@CaraLaitusis
www.linkedin.com/in/Cara-Laitusis
Text Readers for Everyone on All Tests – Getting a Handle on What This Means

Vendor Perspective

Peter Ramsdell, VP Sales
peter@texthlep.com
Assessment objectives to keep in mind

Secure

Consistent

Accurate
Read Aloud Overview

Points of consideration when requesting read aloud for test takers with vision

• Types of read aloud
• Read aloud controls and navigation
• Choices affected by assessment delivery
• External or Internal (embedded) read aloud tool?
• Pronunciation editing
• Practice opportunity
Types of Read Aloud

Human variety

• Staff member reading to students
• Human recorded audio files on CD or folder organized by item
• Human recorded audio files linked to item within digital test
Types of Read Aloud

Digital variety

- Text-to-speech read aloud
- Computer read aloud
- Read aloud with human synthesized voice
- Read aloud with high quality human synthesized voice
- Others
Types of Read Aloud

Visual aspect?

- Read aloud with visual stimulus
- Read aloud with synchronous highlighting
- Read Aloud with dual color synchronous highlighting
- Highlighting by sentence? By word? Or, by sentence and word?

Although the Dutch also had early knowledge of coffee, it does not appear that the Dutch West India Company brought any of it to the first permanent settlement on Manhattan Island (1624). Nor is there any record of coffee in the cargo of the Mayflower (1620), although it included a wooden mortar and pestle, later used to make "coffee powder."
Controls and Navigation

- Speed controls (example: slow, medium, or fast), volume, etc.
- Different combination of colors for the sync’d highlighting
- Navigation through the item:
  - Top to bottom
  - Self selection of sentences and or words
  - Be able to re-read sentences
- Does the vendor need to block read aloud from any areas?
Where will the Read aloud come from?

External 3rd Party Software

Embedded within the assessment
Choices affected by assessment delivery

Locked Down Browser involved?

Cloud Servers

All Cloud, constant access to Internet needed

Web server within district, Internet not needed

Test session delivered to Student Device, then Internet not needed.
Pronunciation Editing

Define for the vendor what is the acceptable accuracy for pronunciations

- Budget for your time and money for a ‘listening’ and adjustment effort to correct pronunciations
- Define what can or can’t be read aloud
- How to handle math equations
Practice Area

• Making a practice area available prior to the assessment is a key for success.

• Applies to both external or internal (embedded) read aloud approaches.
Thank you

Peter Ramsdell, VP Sales
peter@texthlep.com
Text Readers for all: A State Perspective
Melissa Gholson, Ed.D.
Text Readers for all: A State Perspective

Melissa Gholson, Ed.D.

6/31/2017
Accessibility Considerations

- Text to speech is now the leading accessibility option on all plans.
- Concerns exist over the lack of provision during instruction and/or opportunity to practice TTS outside the current assessment system.
- The monitoring of state test accommodations is a requirement for all students with disabilities.
- No professional development or training exists outside assistive technology (AT) programs within special education.
- District (LEA) provision of instructional supports are not equal in expertise or budget.
- How does TTS impact special subgroups (students with disabilities and ELLs).
- How will we know what the impact will be and for whom?
Text to Speech/Read Aloud

- In WV TTS (P01) now surpasses read aloud (P02) as the most assigned support for students with a plan.
- WV recognizes IEP, 504, ELL and SAT plan types.
- Student assistance plans require at a minimum an educator, administrator and a parent (Policy 2510).

<table>
<thead>
<tr>
<th>Code</th>
<th>Total N Count</th>
<th>Enrollment Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>P01-TTS</td>
<td>21704</td>
<td>10.96</td>
</tr>
<tr>
<td>P13-TTS w/passages</td>
<td>3444</td>
<td>1.74</td>
</tr>
<tr>
<td>P02-Read aloud</td>
<td>10427</td>
<td>5.27</td>
</tr>
<tr>
<td>P14-Read aloud w/ passages</td>
<td>2702</td>
<td>1.36</td>
</tr>
</tbody>
</table>
### Universal Text to Speech (TTS)?

<table>
<thead>
<tr>
<th>Supports &amp; Accommodations</th>
<th>Plan Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code</td>
<td>Total N Count</td>
</tr>
<tr>
<td>P01-TTS</td>
<td>21704</td>
</tr>
<tr>
<td>P13-TTS w passages</td>
<td>3444</td>
</tr>
<tr>
<td>P02-Read Aloud</td>
<td>10427</td>
</tr>
<tr>
<td>P14-Read Aloud with Passages</td>
<td>2702</td>
</tr>
</tbody>
</table>
Test Administration

• Universal text to speech may be easier from a test administration perspective.
• Uploading embedded accessibility from student plans requires a high level of technical expertise, provision and monitoring.
• Text to speech is the leading cause of most test administration breaches is supports/ accommodations.
• Universal Text to Speech (TTS) would reduce test resets when support is recent added to an updated plan.
• Differences across platforms.
• Technology issues-voice pack, updates, assuring TTS is working and or available.
Legislation & Legal Issues

- IDEA and assistive technology requirements for students with IEPs.
- IEP teams struggle in provision of FAPE for students that require AT services or devices. Professional development of AT including the provision of TTS is critical for success for teachers, administrators and staff.
- Petcu, Yell & Fletcher (2014) found school districts lost cases by failing to provide:
  - AT assessments-how do we determine if a student needs TTS? How do we evaluate need and whether this provides access or confounds performance? Which functions within the software program will provide access? Which might boost or inhibit performance?
  - Address AT needs-TTS and the need to develop auditory cue strategies and other individual needs such motor or processing.
  - How is the provision of the AT devices or services specified in a student's IEP?- How will districts provide instructional software and training to provide access equally for all learners?
  - Properly implement AT services- making a tool available and not evaluating that the support provides access and does not inhibit performance is tied to the provision of support services. What process will be used to evaluate if additional services are provided and effective? How do we provide the needed professional development across settings and
Auditory & Cognitive Processing

- Auditory processing disorder is a condition that makes it hard to recognize subtle differences between sounds in words. It affects ability to process information.
- Auditory processing disorder includes auditory discrimination, auditory figure-ground discrimination, auditory memory and auditory sequencing.
- Range of listening and learning deficits associated with APD (Musiek & Chermak, 2013).
- Central Auditory Processing Disorders impact: working memory, executive attention, processing speed and alerting attention all of which are correlated with poor literacy and numeracy skills (Ahmed et al, 2014).
- Many children with APD go undiagnosed and the field needs alternate approaches for diagnosis.
Other Considerations

• TTS quality and the need to recognize which kinds of TTS provide access for which learners. (Hearing impaired and parity?)

• Impact on online systems and other technology issues (e.g. not in TIDE, Julie Voice pack was not downloaded)

• Headphones and other technology costs

• Presumption that all learners will be able to determine the settings that work for them (voice, speed, automatic, click as needed, impact on fine motor skills).

• Fidelity of use and failure to monitor usage

• Lack of instructional use and opportunity to practice

• Professional development on supporting auditory remediation

• Policy development

• Studies of impact on student efficacy, performance, by content (math only) or approach for ELA contents?

• Does this impact reading and provision or delivery of passages?

• How does this impact validity of the assessment design or reliability of results?

• Impact on college entrance exams
Questions?

Melissa Gholson
mgholson@k12.wv.us
References


Facilitated Breakout Discussions

- **General education students**
  - Facilitator: Sheryl Lazarus; Notetaker: Kathy Strunk

- **English learners**
  - Facilitator: Martha Thurlow; Notetaker: Linda Goldstone

- **Students with IEPs or 504 plans**
  - Facilitator: Sandra Warren; Notetaker: Michael Moore
Facilitated Breakout Discussion Questions

- What research is needed for about text to speech?
- What common terminology and definitions should be used to more precisely describe text to speech so that they can be used in RFPs?
- What are the implementation and policy challenges that need to be addressed?
Debrief and Next Steps