

WIDA TAC Meeting Notes May 1 and 2, 2017

1. TOPIC A: Comparability (400) and (401) upcoming comparability

- Discuss comparability issues in Speaking and Writing;
 - How do we adjust discrepancies of Speaking and Writing in 402?

Are there comparability differences outside the SEM (Lyle)? If so, what is the cause of the difference? Construct relevant or irrelevant variances? The first step is to identify the source(s) of differences, e.g., are the population differences?

TAC Recommendation: Study where you think the most salient differences are between paper and online, e.g., error in scores, mode differences, population differences, construct differences.

How many cases would a classification decision be effected by the correction? Carol argues that instead of consistency of scores, accuracy of scores.

Look at the correlations between online and paper writing and speaking performances to see if it is performing as expected. Carol M, "Is it possible to have DRC raters watch how local raters score the speaking tests?" Lyle, also suggested looking at speech samples to see if they're systematically different linguistically. Another method might be looking a facets model to look at differences. Another alternative is to conduct a counterbalanced design study having ELs take both forms of the test. Also examine effectiveness of coaching of teacher.

2. TOPIC B: ACCESS 2.0 Psychometric Policies for the following year(s)

- Verification of Speaking and Writing (402 administration) - CAL Shu Jing

Shu Jing outlined content admin plan for series 401. What to do with paper "Static Form 1," for calibrating speaking and writing in 2017-2018. The plan is to not do live calibration with paper speaking and writing.

TAC Recommendation: Look at the early return samples to see if domain scores look like the census data set to see if distribution.

Lyle, is the population changing? Are the item tasks changing? If not, why recalibrate? Carol, M: Will item exposure be an issue with static forms? This (two static forms) is possibly something of concern.

TAC Recommendation: Look at item exposure rates on paper forms to see if this is an issue.

WIDA TAC Meeting Notes May 1 and 2, 2017

- Adjusting Paper and Online score tables;
 - What are your comments or concerns about adjusting ACCESS scores between paper and online modes?
 - What would you recommend that we consider when creating adjusted scores?
 - What should we be looking out for?
 - How would you recommend we communicate this process to WIDA states?

Aki - Equivalent groups are important assumptions to make. Keira mentioned the cross cluster relationship between 302 and 401. Carol M argued that they didn't believe that WIDA could really defend the analysis without more information. She would really like to think a facets introducing the rater or other important facet effects. Lyle, this should only be a onetime thing.

The TAC commented that the current information is not a sufficient basis to make an adjustment. They would recommend waiting for more information, e.g., facets study, rater comparability, construct comparability.

The TAC urged WIDA to state publically that multi-mode assessment within state not be sanctioned.

- Field testing (Embedded Writing test in 2020):
 - Can We Reduce One Scored Task per Form?

CAL shared their PPT on the embedding writing FT. Lyle, suggested doing a d-study and looking at task variance to get a sense for what reduction in tasks would be.

The TAC suggests looking at classification characteristics of cut points, possibly at the Phase 1 cut or PL 4.0.

Regarding the imbedded W FT design, it is important to counterbalance A/BC tasks in the linking slot

Tim, suggested not hiding the last folder position, telling administrators that that position also serves as a linking position and is very important to the test to maintain the scale.

Vince, said FL doesn't communicate FL slots. He would dissuade that.

TAC Recommendation: Communicate to test users that linking/FT items exist on the assessment and why they're there but do not disclose where they are.

- Removing tier cap
 - Do we need to remove tier capping in tier A and B?

WIDA TAC Meeting Notes May 1 and 2, 2017

- Additionally, what difference would tier B and C be after removing tier capping?
- What would the ultimate form Paper be? One form including all tiers, or tier A and C?

Jamal's concern is that removing the tier cap would change the distribution.

The TAC had no objection to removing the tier CAPs.

The TAC had no objection to moving to a two tier system instead of a three tier system.

- Future of Paper forms
 - Does the TAC have comments and/or concerns about either option?
 - Are there any other options we have not considered?

Lyle, for Option 1 do you resolved the comparability issue? I said, likely no. He responded, isn't that a Con for Option 1. I said yes. Lyle's comment, Option 1 seems to be a problem with a hybridized test.

Carol and Lyle, "you need to think about paper vis-a-vis the IEP students taking." Jamal also commented, "do you think the online vs paper accommodations (same type) would provide similar information? I said, likely not."

Lyle, what is the timeframe to do this or get a decision? I said we're considering this but not moving right away.

Another option would be to have a set of specifications that migrates content to an online and paper direction. Jamal's option would be to support the same blueprint for both.

TAC recommendation (Option 3): Explore the use common specifications between online and paper. When in final stages of test development fashion test content to either to online or paper. Then develop folders FT folders unique to the mode. This creates common content across paper and online forms.

- Less than four domain composite scores
 - Do you have questions or concerns about the approach we took examining alternate composites?
 - Do you have other ideas about creating alternate composites?
 - Should there be a limited set of composites that we create?

TACs position would be to provide the procedures to conduct this analysis. Carol, suggested looking at a couple of states to see what the output would look like. Lyle, suggested looking at a CFA to examine the relationship. Correlation between domains is another method suggested by Dorry and Aki.

WIDA TAC Meeting Notes May 1 and 2, 2017

TAC's recommendation: Examine non-four domain composite models and report to them in November.

Dave suggested not doing it, expressing his org would have legal concerns about providing procedures for doing the analysis. Thanks for the assignment. I introduced the issue of optimal weighting. Some discussion about optimal vs effective weighting.

3. TOPIC C: ACCESS 2.0 Annual Technical Report (400 administration)

- Review of ACCESS 2.0 400 administration annual technical report: Paper and Online

Can we give the TAC more guidance on the review of the ATR? Please send a Word version of the draft.

Carol, who are we thinking about the audience? Are psychometricians the only audience? Is there a bigger audience? Think about mental measurements and tests-in-print, technical experts and practitioners?

Lyle wasn't aware that we have an interpretive guide. Carol, concern that the technical report isn't addressing the AUA for practitioners.

Aki, what are all of the materials made available to practitioners to support the test's use, the audience and then look at the ATR to see how that fits and what organization would be appropriate.

Much discussion about how to organize the ATR and possibly the Year in Review report. Lyle commented that this is too much. Carol pushed for us to create the document that meets a broader audience.

Lyle, offered to work with us in framing out an AUA material that could be provided to practitioners. Shu Jing mentioned the Smarter Balanced tech report and its structure.

TAC recommendation - Convert domain CSEM values to reliability estimates at each cut value (2.0, 3.0, etc.).

Carol, there was some confusion about how the AUA and question heading organizational structural. Also, there was a disconnect between how the standards were operationalized in the knowledge, skills and abilities of what is measured. She also had comments about the description of DRC's role vis-a-vis raters.

Jamal mentioned the accuracy and rater consistency of the paper, speaking test. Suggested looking to the audience in a focus or survey about ACCESS test materials' use.

WIDA TAC Meeting Notes May 1 and 2, 2017

Carol, commented on G1WA issue and what QC steps in the ATR. She struggled with the explanation of the table being divorced from the table itself.

Lyle, you have tables that don't have description/interpretation of what the tables are for.

Way forward - Meet with Lyle and frame AUA, identify the documents and audiences that we need to communicate to, survey and focus group (we mentioned four documents to create: Interpretive guide, general uses and interpretive guide, ATR, Year-in-review), connect back with Lyle the week of May 22nd.

Carol, what if you take the AUA out of the ATR and provide two different timelines, e.g., ATR & AUA document.

Vince suggested creating CSEM graphs in addition to the TIF curves.

4. TOPIC D: Planned or Reporting on Validity Studies

- Brief on Validation studies' progress

TAC had comments on the screener and accommodations study.

Screener predictive validity study to include other instrumental variables and other latent variables.

Jamal asked about whether the accommodation study will be looking at the actual IEP. Jamal and Lyle suggested we follow up with schools to double check whether the appropriate. Distribution of scores by accommodation type. Look at time taken for accommodation. Aki, mentioned looking at DIF on accommodations (not sure if we can do that); Compare distribution of item performance of matched samples for students receiving.

5. TOPIC F: Accountability & Peer Review Issues

- Discuss upcoming accountability

No comments. WIDA shared what it knew.

6. TOPIC G: Fall TAC agenda

- Brainstorm topics to discuss in the Fall TAC meeting

Spring Meeting

- Planned or reporting on validity studies

Fall Meeting

- Annual administration debrief including the Year in Review Report

WIDA TAC Meeting Notes May 1 and 2, 2017

- | | |
|---|---|
| <ul style="list-style-type: none">• Accountability & Peer Review issues• Psychometric policies for the following year(s)• AOB | <ul style="list-style-type: none">• Planned research, innovations and/or new assessments <p>AOB</p> |
|---|---|

Add to the fall agenda and updated review of the AUA and other relevant documents. Preliminarily identify (based on AUA structure) where validity studies are missing. This preliminary AUA documents and validity study possibilities should be delivered by beginning of October. Carol asked about other documents provided as well. Dorry mentioned the need to provide the totality of materials WIDA provides for assessment use and interpretations. TAC requested that WIDA provide an overview of all the materials is uses to support ACCESS score use and interpretations.