

**Nebraska Technical Advisory Committee Meeting**  
**Nebraska Department of Education**  
**March 2, 2023**  
**9:00 – 1:00**

**9:00 a.m. – 9:15 a.m. Welcome and Introductions**

**Present:** Chad Buckendahl, Josh Fields, Christy Hovanetz, Jeff Nellhaus, Linda Poole

**Approval of Minutes from 1/12/23 and 1/27/23:** After discussion noted below, Chad called for a motion to approve minutes as presented: Josh Fields made the motion, Linda Poole *seconded*. Motion approved, 4-0. Christy Hovanetz abstained.

**9:15 a.m. – 10:00 a.m. NWEA/ACS Ventures presents standard setting procedures for summer 2023**

**NWEA/ACS Ventures presents standard setting procedures for summer 2023**

Prior to the presentation, a TAC member noted his company is working with NWEA to conduct the ELA and math standard setting planned for the summer of 2023. A colleague, with ACS Ventures, is presenting the standard setting plan for input from the TAC.

**Questions:**

- **How much do we want to deviate for the math to address specific concerns or should the same processes/procedures be used for both ELA and math to counter any concerns later?**  
**TAC:** What is the criteria for ordering items for the ordered item booklet? Using P value, IRT statistic? NWEA used the IRT difficulty to order the items. ACS Ventures proceeded to describe the process for both the ELA and Math since the review items are different (i.e., ELA has passages). Would you want to bring all subjects to the policy review from the AAAC (Assessment & Accountability Advisory Committee)? Yes. **TAC:** Often panelists take the test. You should consider including this activity. Point around rounds, just 2; what is the nature of the feedback after the first round? **ACS:** Nature of the feedback is making sure what cut scores they set and how it translates to a student's score, and how it relates to what their colleague did. Also show impact of cut scores on results. **TAC:** It will be useful to provide impact data but consider rather than providing all feedback after the 1<sup>st</sup> round, do so after the 2<sup>nd</sup> round and provide judgments.
- **What external sources of information should be provided to the panelists to consider when evaluating the standard setting results (after Round 1 or Round 2)?**
  - o **2022 state-level NAEP results**
  - o **NWEA comparison of "proficient" cut scores by RIT values across states**
  - o **Others?**

What additional information does TAC suggest? NAEP (in second round of impact data) and NWEA RIT scores make sense. Be careful comparing data state to state; it would probably not be good to use with standard setting team. Performance Level Descriptors (PLDs) should drive the cut scores. Can look at NAEP PLD vs. Nebraska PLD for similarities and contrasts and general definitions of PLs are consistent.
- **Would it be helpful to do a vertical articulation review after the first set of cuts are set?**
- **TAC:** Does not strongly recommend this. You want content people to focus on anchoring the content of the PLDs to the assessments and cut scores.
- **What other suggestions does the TAC have for the design or execution of this standard setting?**

## **What suggestions does the TAC have around Science Standard setting processes?**

- **TAC:** Does giving them the yes/no from the previous year compel the panel to agree? May want to bring in external data like NAEP to compare proficiency.

**TAC:** When doing the intro, we hope that the conversation is around student success and what a successful student can do and what is meaningful to students – if individual students can do this will they be successful in next grade. As come back at showing impact data, really focus on student not on just the values and data.

## **10:00 a.m. – 11:00 a.m. DRC presents their evidence of validity for the alternate assessments**

### **Are there additional things that we should be documenting or incorporating into our process?**

**TAC:** Have you done cognitive labs to see how the students are engaging in the process? It not required but maybe with particular item types or with new things it is part of the field tests especially with any new kinds of measurement of students.

**TAC:** With content validity in terms of how blueprint is developed, what sort of process occurs to define what is included and what isn't in terms of that connection to the general content standards compared to the extensions. It is important to explain rationale as to why certain areas of the content are prioritized over others.

## **11:00 a.m. – 11:15 a.m. Break**

## **11:15 a.m. – 1:00 p.m. DRC concludes the evidence of validity presentation**

### **Is there additional evidence that DRC should consider presenting for peer review?**

**TAC:** For ELA at grade 11 range suggested 19 out of 25 – in terms of fit, is there anything that can speak to regarding why some of those might not have fallen in range. It would be helpful as develop more items to understand what causing it. Be good to know at grade 3 and 11 if sample characteristics or distributional characteristics within population are influencing. With these models, you're not supposed to be sample dependent but when dealing with 200 students per grade, it may influence the observed parameter estimates and the diagnostics as a result.

In terms of correlations shared with relationship to strand level with English to English, English to math, etc. We're surprised that they are that closely correlated based on diversity of population. Is it because the kids are bottoming out? Might make sense to look at the reasons why. It is worth exploring to know that some of it is related to the way that disintenuation works given population characteristics. Sometimes there's higher correlation between English and Science than two of English strands. Want to make sure measuring unique things between content area and one skill isn't the dominate skill.

In dimensionality checks do you include ELA, math, and science all in same dimensionality? Within grade or subject if put all of this into a PCA would you get 3 distinct dimensions (Math, ELA, Science) or is there a single underlying dimension that explains all of them? Correlations among three subject areas need to be explored to ensure measuring distinctly from one content and another.

May need more qualitative data around how assessment is created on the front end. How development of extensions creates this access for the diversity of students. Are there sufficient access points for students to participate in the assessment? The purpose of this assessment is that we know they can't access general so giving them bridges to demonstrate that instruction is showing up in progress. How

does alternate connect to general content standards? Demonstrate how instructional supports connect and provide opportunities. Show how you are providing supports and PD to ensure that they have every opportunity to learn and then demonstrate on the assessment.

### **Final thoughts/questions**

Is there an adaptive option with this extremely diverse population? A staged adaptive approach would be good and some of the variability and high item total discrimination correlations across grades illustrates the range of ability in this population. It may be good to look at states who have staged adaptive tests such as Indiana and Michigan.

**Prior to approval of minutes**, TAC requested clarification regarding the reasons for the standard setting in 2023 and 2024. As well as what NSCAS Growth as a through-year/course assessment implementation. Director Clark provided the description of NSCAS Growth. Director Clark reviewed the concerns regarding the math standard setting process in 2018 that led to a determination to conduct a full standard setting for the math legacy standards. For 2024, standard setting for new math standards will depend on significance of change between legacy and new math standards. Input from TAC will be elicited on the significance of change.

**Next TAC meeting is scheduled for 9 a.m. – 1 p.m. on May 11, 2023.**

### **Topics for future TAC Meetings:**

- Evidence of validity for ACT and Cambium for ELPA21
- Accountability around chronic absenteeism and ELPA21 cut scores for accountability
- Test design around the concept of through year