

**Governor's Technical Advisory Committee Meeting  
Nebraska Department of Education  
February 12 and 13, 2018**

**Cornhusker Hotel, Lincoln, NE**

**Day 1**

**8:30 am-4:30 pm**

**February 12, 2018**

**8:00-8:30**      Check-in- Breakfast

**8:30-8:40**      **Welcome and Introductions- Valorie**

**8:40**            **Approve Minutes- Chair- Brian Gong**

*Document 1-Minutes Technical Advisory Committee October 31, 2017*

**8:40-9:10**      **Preface- Explanation of Nebraska Student-Centered Assessment System**

NDE has revised its assessment system from the former NeSA system to the Nebraska Student-Centered Assessment System (NSCAS).

*Document 2-Overview NSCAS System*

**9:10-10:10**    **NSCAS Summative General- Grades 3-8 Standard Setting Design and ALDs**

NDE will be engaged in standard setting and standards validation activities in 2017–2018. The mathematics standard setting plan using the Item-Descriptor (ID) Matching method and ELA cut score review are presented in Documents 3-4. The mathematics achievement level descriptors (ALDs) are presented in Document 5, and the ELA ALDs are presented in Document 6.

*Document 3: Mathematics ID Matching Standard Setting Workshop*

1. NWEA revised the panel composition of the standard setting plan based on the TAC's recommendations. Does the TAC have any additional feedback on the ID Matching standard setting workshop plan?
2. May NWEA move forward with this approach?

*Document4: ELA Cut Score Review*

1. Does the TAC have feedback on the number of items around the target cut score panelists should review, defined as the threshold region?

2. Because standard setting occurred last year, NWEA anticipates pre-matching the easiest items in the OIB and the hardest items in the OIB to focus panelists on expert matches for the threshold region. Do you anticipate any issues with this approach?
3. Is it necessary to conduct three rounds of judgments during a cut score review, or is messaging last year's standard setting as Round 1 sufficient?

*Document 5a- Math ALD Workshop Plan*

*Documents 5 and 6: Achievement Level Descriptors for Mathematics and ELA*

1. Does the TAC have feedback on the Range ALDs? Do the ALDs:
  - Describe what students can do?
  - Define differences in content across performance levels rather than the frequency (e.g., partial, minimal) with which students respond to content?
  - Describe the contextual or scaffolding characteristics needed so a student can demonstrate the skill?
  - Increase in cognitive processing complexity across levels in a cogent way?
  - Provide a mental picture of increases in skill across levels?
  - Describe the increasingly more sophisticated ways of reasoning in the content domain that follow one another as a student learns and grows in their skill?

**10:10-10:20 Break**

**10:20-11:50 Continue NSCAS Summative General- Grades 3-8 Standard Setting Design and ALDs**

**11:50-12:35 Working Lunch (Positive messaging regarding transition to CCR)**

**12:35-2:35 Computer Adaptive Testing Simulations (NWEA)**

*Documents 7a and 7b: CAT Simulations*

Information regarding the technical adequacy of computer adaptive test (CAT) events and the item pool is important for building an argument regarding the interpretation and use of test scores in relation to the standards and test construct. The simulations are intended to provide information regarding the degree to which students receive comparable representations of content. Therefore, the simulation study serves two main purposes. First, it allows the

state to determine if the item pool is sufficient to find a feasible set of items for students within a grade who have different ability levels. Second, it evaluates the functioning of the engine's item selection algorithm to ensure that the state's construct (blueprint) is being represented as intended. NWEA reports the content constraints that are supportable, including the depth of the pool to measure strand, standard, depth of knowledge (DOK), and point allocation for each strand for 2018. NWEA also reports on the ability of the CAT to estimate student ability with investigations of bias (i.e., any significant variance between true and estimated theta), mean standard error (MSE), root mean square error (RMSE), and the correlation between true and estimated theta.

1. Does the TAC have questions regarding the simulation study NWEA provided?
2. Does the TAC have any suggestions as we replicate this process for 2018–2019?

**2:35-2:45**      **Break**

**2:45-4:15**      **AQuESTT Classification and Priority School Alignment to ESSA plan and CSI/Targeted Improvement School Selection**

Since the establishment of the AQuESTT framework in 2015, a 10-year Strategic Plan was developed and launched by the State Board of Education in December 2016. Additionally, in December 2015, the Every Student Succeeds Act (ESSA) was passed into law mandating all states to submit a consolidated plan focused on ensuring a quality education for all students regardless of race, zip code, language proficiency, or disability. Finally, new college and career readiness standards and assessments have also necessitated changes to the AQuESTT classification system.

These three developments have driven conversations on the new classifications system.

Changes to AQuESTT should happen in tandem with new school improvement tools, updates to accreditation guidelines (Rule 10), and the implementation of ESSA. In sum, the time is ripe for a robust review of Nebraska's accountability system through broad stakeholder engagement, informing the state on the new system, and implementing school improvement strategies with fidelity.

*Document 8: Nebraska ESSA Plan Clips*

**Discussion of “Big Questions”**

*The NDE would like TAC’s opinions and input on big picture changes that will be made to AQuESTT in the next couple of months. This input will be provided to NDE staff and stakeholders who will be making recommendations to the Commissioner on updates and changes to the AQuESTT classification system.*

School Quality or Student Success Indicators:

1) Chronic Absenteeism

- Big Question:* How should chronic absenteeism be integrated into AQuESTT? What options does NDE have when integrating chronic absenteeism into AQuESTT?
- What is the TAC’s opinion on the definition outlined in the ESSA plan?
- What is the TAC’s feedback on Nebraska’s proposed system using 50% reduction as the basis of the calculation?
- What challenges does NDE need to consider?

2) Science Indicator

- Big Question:* How should science be included in the AQuESTT system as a separate indicator?
- What is the TAC’s feedback on the use of proficiency for the science indicator?
- Are there other aspects of science that we could consider when designing the indicator?
- What challenges does NDE need to consider?

3) Progress toward English Language Proficiency

- Big Question:* What are the TAC’s suggestions for inclusion of English language proficiency in AQuESTT?
- How do we measure student-level progress?
- How does student progress roll up to school-level indicators?
- What challenges does NDE need to consider?

System Changes:

4) Indicator Weighting

- Big Question:* The development of our ESSA plan brought up the question of weighting of indicators. After reviewing proposed weighting included in the state’s ESSA plan, what are the TAC’s opinions on the subject?
- Given the current “impact weighting” of our system, what changes need to be made to functionally make weights more growth-oriented?

5) Scale Scores vs. Proficiency

*-Big Question:* Nebraska has moved from the use of scale scores to proficiency in measuring academic achievement. What input would you offer on this proposal?

-What are the TAC's recommendations for accommodating this change in the Status indicator of AQuESTT?

-How will this change affect other indicators in AQuESTT?

#### 6) Stability of Classification

*-Big Question:* The NDE has decided to move to classification of schools annually. One concern has always been on the variability of a school's classification in subsequent years. What suggestions does the TAC have on creating a more stable classification system?

-What are the sources of variability?

-How can variability be smoothed from year-to-year?

#### 7) Comprehensive Support and Improvement and Targeted Support and Improvement

*-Big Question:* After reviewing NDE's criteria for identification, support, and exit from Comprehensive and Targeted Support status, what input would you offer on this proposal?

-What input does the TAC have on the index system proposed for Targeted Support and Improvement school identification?

**4:15-4:30 Winter 2019 TAC Date- Please bring calendars**

**Governor's Technical Advisory Committee Meeting  
Nebraska Department of Education  
February 12 and 13, 2018**

**Cornhusker Hotel, Lincoln, NE  
8:30 am-3:45 pm  
Day 2**

**February 13, 2018**

**8:00-8:30** Check-in- Breakfast

**8:30-8:40** **Welcome-** Comments from Day 1

**8:40-9:50** **NSCAS Summative Peer Review: Test Interpretation Framework (NWEA)**

*Document 9: Principled Test Design Based on ALDs*

The design and validation of an assessment system, intended for both formative and summative purposes, requires careful development processes, especially when such assessments are intended to support interpretations regarding how student learning grows more sophisticated over time (Pellegrino, DiBello, & Goldman, 2016<sup>i</sup>). Under a principled approach to test design based on ALDs, the evidence needed to draw a conclusion about where a student is in their learning of content is made explicit in the ALDs and items are developed specifically to those evidence pieces (Huff, Warner, & Schweid, 2016<sup>ii</sup>; Egan, Schneider, & Ferrara, 2012<sup>iii</sup>; Schneider & Johnson, in press<sup>iv</sup>). Under such a model, increases in cognitive processing complexity (e.g., DOK) are intended to be embedded into evidence statements across achievement levels in a cogent way and to interact with content such that empirically, the features of cognitive processing and content interact to affect item difficulty.

A principled approach to test design is intended to support the validity of inferences about the student's stage of learning and the content validity of the assessment as a measure of student achievement. Under such a score interpretation model, construction of test blueprints should not separately and distinctly treat DOK as a separate constraint. Instead, DOK should be present as a component of evidence that supports interpretations regarding the stage of thinking sophistication the student is at during the time of the test event.

1. Does the TAC have any feedback on the framework to support the state's test score argument for peer review purposes?

**9:50-10:05 Break**

**10:05-11:30 NSCAS-Summative Alternate Assessment of Mathematics (NSCAS-AAM) Grades 3-8 and 11 Standard Setting Plan**

NDE will engage in standard setting for the NSCAS-AAM test (i.e., alternate mathematics) in summer 2018. The proposed standard setting will use a Yes/No Angoff technique, similar to that used in 2017 for NeSA-ELA Alternate. The plan for this work is presented in Document 10.

Document 10 and 10b: NSCAS-AAM Proposed Standard Setting Plan

1. Does the TAC believe this plan is sufficient to establish defensible cut scores for NSCAS-AAM?
2. How should NDE and standard setting participants consider the connections between NSCAS-AAM and the analogous tests of general education mathematics?
3. Should the NDE expect a relationship between the NSCAS-AAM results in grade 11 and the state's ACT results?

**11:30-12:15 Working Lunch (Management of multiple testing vendors—clarity to schools)**

**12:15-1:15 NSCAS Summative General- Grades 3-8-Peer Review: Guidance on Three Research Studies (NWEA)**

NWEA is contracted to provide three research studies in support of the peer review process.

1. Does the TAC have guidance on what studies are most critical for a successful peer review process?
2. What is the current thinking in regard to accommodation research?
3. Are accommodations required to be researched on each testing program or are reviews of findings across the literature sufficient?
4. Should focus be placed on device comparability?

**1:15-1:30 Break**

**1:30-3:00 Continue NSCAS Summative General- Grades 3-8-Peer Review: Guidance on Three Research Studies**

**3:00-3:30 NSCAS Summative Delivery Robustness: District/School Preparedness**

*Document 11: Delivery Plan*

NWEA has an established plan to support District partners in the implementation of NSCAS Summative General. The technical requirements for the summative assessment and the protocols for test administration mirror NWEA MAP Growth, which is used by nearly all districts in Nebraska. Preparedness for new and existing partners has included virtual and face to face trainings, a team of dedicated help desk agents trained to support Nebraska partners, and a network of Technical Consultants to support implementation prior to and during testing. The NWEA delivery plan also includes a comprehensive communication plan tailored to the unique needs of NSCAS.

**3:30-3:45**      **Wrap Up**