

Technical Advisory Committee
Nebraska Department of Education
November 7, 2012
Cornhusker Marriott Hotel Lincoln, NE 8:30 am-3:00 pm

Present: TAC Members: [Brian Gong](#), [Richard Sawyer](#), [Linda Poole](#), [Frank Harwood](#)
DRC/CAL Staff: [Patricia Johnson](#), [Sherri Wolfe](#), [Lucy Liu](#), [Dave Chayer](#), [Linette McJunkin](#), [Janet Hensley](#),
[Stacey Reasoner](#)
NDE Staff: [Jeremy Heneger](#), [Ed Foy](#), [Ted Larson](#), [Valorie Foy](#), [Russ Masco](#), [Freida Lange](#), [Scott Swisher](#),
[Roger Breed \(AM\)](#), [Brian Halsted \(AM\)](#), [Sharon Heater](#), [Diane Stuehmer](#)
Governor's office: [Sheryl Wolff \(AM\)](#)
Consultant: [Bill Auty](#)

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

[Open Meetings Act](#)

8:40 AM **Approve Minutes – Brian**

[Summary was approved](#)

8:40-9:45 **Update: Check for Learning Valorie, Jeremy**

(Document 1)

The group reviewed the Check for Learning [C4L] system, which is functioning well for schools; some comments were offered concerning the quality of items. NDE staff explained the extensive review of items, which includes intensive content review and feedback from users. It was suggested that the assessment department share the review process with school district personnel so they understand that the items are reviewed extensively and that weak items are discarded. (This information was included in the Check for Learning presentations given in late November and early December 2012 throughout Nebraska.)

The group suggested that NDE work with districts to track use of C4L items across time and to develop district expertise in building test forms that produce the information districts want, while at the same time encouraging districts to avoid using C4L to develop high stakes tests. If districts are interested in additional quality over time, NDE could invite districts to partner in research studies, for example, in program evaluation. NDE might also survey districts that use only C4L, use their own interim systems and C4L, or use only their own interim systems.

Brian Gong framed the value of C4L as one step in a three-part system:

- [Diagnosis takes place with the C4L assessment.](#)
- [Prescription provides the description of the intervention needed to improve student learning](#)
- [Treatment is the carrying out of the intervention](#)

[Suggestions were given for clarifying the data charts of results available to districts in C4L, followed by a discussion of the use of C4L results compared to use of NeSA results. The availability of C4L data precludes the tendency of districts to over-interpret the data](#)

available from NeSA testing. NeSA is a summative test while C4L provides richer more robust information on student learning in relation to Nebraska state standards tested on NeSA.

9:45-10:00 Break

10:00-11:30 Update: Nebraska State Accountability [NeSA] 2012 Valorie (Document 2)

The group reviewed and discussed NeSA results. NeSA tests are pre-equated using embedded items and then post-equated; continuous post-equating is recommended. The discussion of NeSA-Writing included the predictable mode of testing at grades 4, 8, and 11 from year to year and the possible limitation that might result in instruction, the relationship of domain scores (“chunking of scores”), and the inclusion of the writing verification process in the writing technical report . In a review of the writing scoring rubric, the group indicated that weighting revision (sentence fluency/conventions) less than ideas (ideas/content) and organization makes sense.

The upcoming standard setting for 4th grade writing was discussed, and the group determined that the methods used for grades 8 and 11 in 2012 were effective, and suggested that the standard setting for 4th grade writing should include the same procedures.

In the discussion of NeSA- Reading, Math, Science, the group suggested review of NeSA results over time and evaluation of the consistency of results, including maintaining equal difficulty across grade levels and possibly checking against an external measure. Brian Gong suggested that school districts need to examine many factors, such as student characteristics and discipline reports, not only the test itself, in order to improve NeSA scores. Discussion included the suggestions that NDE explore the professional development that has grown in districts, based on NeSA results, to determine what professional development raises scores and the ways in which NeSA tests help school districts reach their improvement goals. Richard Sawyer suggested that the research by Steve Sireci on a framework of content validity would be helpful. (I have attached the article.)

The group discussed negative growth in reading from 7th to 8th grade and the flattening out of scores at grade 11, a phenomenon that follows the national high school trend.

11:30-12:30 Working Lunch-- Discussion of NeSA Security Valorie (Document 3)

The group reviewed the security document being prepared for release by NDE and discussed security procedures and possible security forensics. Some information is available through the CAL on-line testing system, such as log-in times and ability to track when a student leaves an item, when the student returns, and what change takes place in the answer. DRC can also provide erasure analysis.

In reviewing the Security Breach form, the group suggested looking for patterns in the reported breaches and checking with other states in regard to the number of incidents.

We do not yet know what constitutes a significant improvement in test scores. Data in subsequent years will inform the interpretation of changes in the data in regard to test security. The group suggested that NDE do a cost/benefit analysis of data forensics. Richard Sawyer indicated that statistical analysis can detect improbable events, but to determine if a security event has happened, a second arm that detects whether the change in data is legitimate or not needs to be in place.

12:30-12:45 Break

12:45-2:45 Update: Nebraska Performance Accountability System [NePAS] Valorie (Document 4)

The group reviewed the first year of the release of Nebraska Performance Accountability System [NePAS] in light of the theory of action framework: If school districts receive low ranks, do scores improve? The group discussed the issues of validity of rankings and reliability of rankings over time and the need for districts to examine growth as related to academics, but also to factors such as population changes. In NeSA scaling, a growth of zero equals one year of growth, but NDE will need to study the data over years. Brian Gong suggested that any changes to NePAS be done for the current 2012-2013 school year, and the system could then be evaluated over the next five years.

The ceiling effect on results in NeSA was discussed; Brian Gong indicated that Nebraska scores have room to grow and that moving students from “Below” to “Meets Standards” occurs more readily than moving students from “Meets” to “Exceeds.” Some states are using AP courses to push rigor in high school, and concepts of career-ready versus college-ready was discussed in relation to math standards. NDE staff indicated that an ACT pilot is being conducted in Nebraska with eight schools. The group suggested that NeSA items be shared with parents; the NeSA practice tests are available on-line to the public, so parents can access them.

Richard Sawyer commented on the interpretation of improvement scores: what is the extent by which improved scores can be an indication that the improvement is not by chance. Most observed changes are not statistically significant (such as fluctuation in ACT scores in larger schools). Since NePAS includes rankings for all groups, some as small as 10, NDE might consider including confidence levels in the on-line report (State of the Schools Report). We can look at the variability of indicators over time and track stability of rankings, changes due to standard deviation, and changes due to chance. Rankings that have great fluctuation over time will be taken less seriously by districts and stakeholders. If NDE wants the rankings to be considered seriously by districts, the rankings cannot vary erratically from year to year so as to be meaningless, especially considering the small size of some of the groups that were included.

The group discussed breaking the rankings into stanines (see NAEP) and the advisability of reducing the number of rankings. With the multiple rankings the interpretation is up to the district. The more the state collapses the rankings, the more the district story becomes implied; single rankings cause the media to focus on the rankings even more than they currently do. Brian Gong stated, "No other state has implemented a system similar to NePAS so there is not a research model to support or decry what you are doing."

2:45-3:00

Wrap up and next steps.

In summary, the TAC discussed, recommended, and acted on all addenda items
The TAC:

- Reviewed and was supportive of the use of the interim system C4L.
- Suggested extending the value of C4L by encouraging districts to use C4L to diagnose learning problems and develop their own interventions.
- Reviewed Nebraska NeSA scores in Writing, Reading, Math, and Science for the years available.
- Emphasized extending the use of NeSA Reading, Math, and Science data to promote professional development to improve scores.
- Recommended including the NeSA-W verification process and results in the Writing Technical Manual.
- Recommended research on the variability of domain scores on NeSA-W.
- Recommended the procedures used for 8th and 11th grade standard setting in writing be used for 4th grade standard setting.
- Suggested that NDE review scores to determine factors that affect scores, but need five to seven years of data.
- Recommended continuous post-equating of NeSA tests.
- Recommended that the cost/benefit of a security pilot be determined.
- Recommended looking for patterns in the reported security breaches.
- Recommended checking with other states in regard to the number of security incidents.
- Reviewed NePAS system and the results for school districts.
- Suggested that any changes needed to NePAS take place in the current year, to be used for five to seven years to analyze the results.

Dates for future TAC meetings

- Spring 2013- April 24, 2013
- Fall 2013- December 4, 2013