

Technical Advisory Committee – Summary Notes
Friday June 25, 2010
Cornhusker Marriott Hotel
8:30-3:30 PM

Present: Brian Gong, Dallas Watkins, Richard Sawyer, John Poggio, Bob Evnen, Patricia Johnson, Ron Mead, Jessica Tickle, Roger Breed (morning), Scott Swisher (afternoon), Cheryl Wolff (afternoon), Marilyn Peterson, John Moon, Ed Foy, Carla Osberg, Carol McClain, Donlynn Rice, Ted Larson, Jan Hoegh, Freida Lange, Pat Roschewski

8:30 AM Welcome and Introductions – Pat

8:45 AM Approve Minutes (Document 01) – Brian

8:55-10:00 Update: NeSA-Reading Operational Report (Document 02) – DRC (Patricia and Ron)
NeSA-Math Field Test Report (Document 03) – DRC (Patricia and Ron)

DRC reviewed the documentation and discussion was focused on the following topics:

- Differences in testing mode, differences in numbers and reasons why
- Online testing times – the average testing time per grade was presented. Richard suggested that CAL capture/share the distribution of time – the longest testing time and the shortest testing time. This is information that could be provided to districts, perhaps helping them with their scheduling.
- Dallas talked about the time spent testing. The transition between the old local system and NeSA means that districts are doing some of both, and teachers feel they are spending too much time testing. Educators in the field are also fearful of what is ahead – the Common Core standards and assessment.
- Brian suggested that NDE/DRC investigate whether the amount of time required might be reduced between test administration and the issuing of reports.
- Talked about the length of NeSA compared to other state tests. Jessica indicated ours was comparable but on the shorter side of some tests. Ron indicated that the number of NeSA items were fewer but overall student-level reliabilities are fine.
- John introduced the notion of computer adaptive testing as a means to save time, but that model requires a huge item bank which Nebraska does not have yet, but that we might consider collaboration with other states.

- Ron indicated that our average p-value on the field test was .70, meaning that on average 7 out of 10 students were answering items correctly. The fact that this is a standards-based test, the p value level is not concerning, but the real issue will be where teachers set the proficiency levels during the standard setting.

10:00-10:15 Update: Standard Setting – (Document 05) Pat and DRC

In preparation for the Bookmark standard setting scheduled for June 28-30, the following topics were discussed:

- Ron indicated that the Contrasting Group Method – conducted in March included on average 59 teachers per grade level, (413 teachers projecting levels of 8,831 students) and occurred prior to testing. The results of the CG method will be shared during the Bookmark Standard Setting. The results, Ron reported, were consistent across grades.
- During the Bookmark Standard Setting, both NAEP and ACT data will be shared with the panelists. Over 100 teachers will be serving as panelists in June, and this will include those teachers who are panelists for the alternate assessment.
- Brian suggested that the detailed PLDs for both NeSA-R and those for NAEP be shared with the panelists and recommended that panelists be reminded that NAEP’s four proficiency levels are different from Nebraska’s three levels, although they share some of the same labels.
- At the July State Board Meeting, the results of the methods will be shared for the board approval of the cut scores. The board will receive an appropriate range of cut scores, the recommended cut scores, and the supporting data. Bob asked about the proficiency labels and whether or not they could be changed. Patricia indicated she would check on that. Pat asked for her to check on the cost.
- Richard reminded the group that the supporting summary data need to be included in the NeSA-R technical reports.

10:15-10:30 Break

10:30-11:30 Long Range Assessment Schedule (Document 06) – DRC – Jessica Tickle

Jessica shared the long-range assessment schedule with the group, and discussion followed:

- Brian suggested that a more explicit field test plan be built into the long-range planning document. He suggested that the chart be adjusted to show the needed field test items to match the blueprint and also include the actual yield of the field tested items so that the plan reflects an active, targeted management of the field-testing/development effort to yield operational forms.
- John noted the risk of having only one operational form, but in order to have more, more items need to be developed and field tested.
- Need to have conversations about how to manage and schedule future item development, where to prioritize, where to target funds, where items are most needed, and what plans are possible to meet future goals with an item bank.

11:30-12:15 Working lunch

Update: Reporting NeSA Results (Documents 07,08,09) – Pat

During lunch, Pat shared the NeSA individual student report and the display of results for the State of the Schools Report and pointed out the changes that had been made to the reports based on an earlier TAC meeting. A discussion followed about adding the Student Percentile Rank to the student report. Brian indicated that although in the past it might have been confusing to have a percentile rank on a criterion-referenced test report, he believes that parents would understand a percentile rank and it should be added. Richard concurred.

Patricia will be checking on the feasibility and cost of adding the student percentile rank at this stage of the process since the reports have been designed, and the NDE has signed off on them. Patricia will get back to the NDE with the feasibility and cost of doing that and still meeting the August 16th release timeline.

Brian suggested that if the committee had comments about the Interpretive Guide to email Pat.

12:15-1:00 Comparability Study (Document 10) - DRC/CAL

John outlined the plan for the comparability study:

- The group indicated that even though Nebraska has a large percentage of students tested online, the need still exists for a comparability study in Nebraska.
- Richard suggested that rather than the process of “test construct analysis” that John conduct a mode comparison study of “like groups,” such as special education or ELL students that took both modes, etc.
- Richard suggested that comparability needs to be addressed in the design phase and perhaps a panel of experts be convened to address the comparability of the mode forms.
- Brian pointed out that the NDE should be attentive to validity issues and to be sure that there are no changes in the construct between the two modes (scrolling and chunking with online, for example, that is not readily available in paper/pencil).
- Brian and Richard asked John to reconsider the comparability plan and the revised plan should be submitted to the NDE and the TAC.

1:00-2:15 PM State Accountability – Next Steps – Pat and Bill Auty (by phone)

Bill Auty and Pat overviewed a timeline and pilot study plan for the accountability discussion:

- At this time a board subcommittee is reviewing its options for building an accountability model. In response to Brian’s questions, Bill and Pat confirmed that the board is beginning with a public report of accountability vs one of sanctions and rewards.
- The Board wants to have a proposal for an accountability model in place by September. They will be having further discussion about the purposes, the design, and indicators to include in future meetings.
- The plan, under development can include growth by 2012, but the first step is in getting the first information publically displayed in as clear a way as possible.
- Brian suggested the need for some “interpretative language” of the accountability system.
- The topic of performance indicators was raised – what should be included in the accountability of schools? Richard pointed out that some factors are out of the district’s control. Dallas indicated, for example, that attendance might be one of those factors. Dallas also mentioned that small numbers, or as Brian rephrased it, “units of reporting analysis” are of concern.
- In answer to Bob’s question, Bill indicated at this point “growth” is not a requirement, but it might likely be in future years. Brian indicated that the conversation about growth must begin now to set the baseline for the future,

and then examine multiple growth models to see which matches the Board's purpose. Hence, the Board all need to agree on the purpose of the model.

- **Bill is conducting a pilot study on one growth model, the student growth percentiles model, currently being used in Colorado and being adopted by many states in the country. He is using two districts' data. The reason those two districts were selected is that they both have "summative" tests scores used for local assessment reporting rather than classroom-based assessments used throughout the year.**
- **Brian reaffirmed that the State Board subcommittee must first establish its purposes so that the development of a model can evolve.**

2:15-2:45

Break – DRC/CAL representatives were excused from the room while two additional agenda items were discussed:

Statewide writing – Pat briefed the TAC on the issues that are initiating a transition plan in statewide writing. Input is being sought from the TAC and will be sought from the State Board members at the July State Board meeting:

- The revised reading/writing standards place greater emphasis on conventions and other writing traits; therefore the alignment of the current writing is being reviewed.
- For the last 10 years the writing assessment has been scored holistically. Analytic scoring would provide trait-based scores and more information to teachers.
- Some school districts as well as business and industry groups are pushing for online writing assessment as a more "college and career ready" approach to writing.
- The legislation requires the assessment of only one grade level in writing; since 2003 the NDE has been assessing three grades, and budgetary issues are causing a review of current assessment processes.
- The current contractor indicated they have facilitated scoring for ten years, and the NDE should explore other options.

It is important that NDE provides a transition to writing and not force abrupt change for schools, so the department is open to the TAC and the State Board's suggestions about the planning. The NDE has tried always to transition into change, and is suggesting that for the 2010-2011 year, grades 4 and 8 be continued as a paper/pencil assessment with holistic scoring. The NDE recommends that the 11th grade begin a "pilot" transition to online writing administration and analytic scoring. The transition would continue over the next 3-5 years with other grades moving to analytic scoring and online test administration.

A request for proposals was posted for writing contractors on May 31st, and five vendors submitted plans to assist NDE with the transition. The decision of a contractor must be made so that a contract can be awarded no later than September.

The TAC offered the following advice and information:

- **Be certain that the prompt initiates the intended outcomes.**
- **Some states score mechanics objectively and separate from the writing prompt.**
- **Scoring cost is often a factor in using analytic scoring versus holistic scoring, as well as whether the analytic scoring dimensions are highly inter-correlated.**
- **It might be that NDE will want to have a better writing model in one grade rather than trying to do three grades.**
- **Think through the issues of transitioning from one vendor to another.**

2:45-3:15 Discussion: Common Core Assessment (Document 11)

Brian and Pat described the two consortia that have submitted applications to the USDE for the Common Core Assessment. One, called “PARCC” was managed in the proposal stage by ACHIEVE, Florida is the fiscal agent, and has 27 states as members. The other, SMARTER BALANCED” is sponsored by a group of 31 states with Washington State the fiscal agent, and WestEd as the proposal’s managing partner. (A third consortium, SCOBES, that is focusing on developing “end-of-course board examination-like systems” applied under a different portion of the RTTT and is less relevant to Nebraska.) Both consortia have submitted proposals to USED. Essentially the differences between the two consortia are the following:

PARCC is computer delivered assessment; SMARTER BALANCED is computer-adaptive assessment, adjusting to the student’s level

PARCC has less teacher involvement in development and scoring than SMARTER BALANCED.

SMARTER BALANCED has a greater emphasis on formative and interim/benchmark assessment.

In the discussion, the TAC advice included the following points:

- **While planning the Nebraska testing process, consider the implications in preparation of 2014-15 when the Common Core Assessment is scheduled to be administered.**
- **If Nebraska adopts the Common Core Standards, be sure that teachers are familiar with them a couple of years before the Common Core Assessment.**
- **Monitor the costs of the CCore assessment.**

3:15-3:30

Wrap up and next steps – Brian and Pat

Summarizing comments and recommendations:

- ***The TAC was supportive of the development and direction of the NeSA testing process.***
- ***The TAC has recommended a revised procedure for the comparability study. John Poggio will be in touch with Brian and Richard regarding a revised approach.***
- ***The TAC has recommended that the student percentile rank be placed in the student report of results.***

Next Scheduled Meeting: Tuesday October 19th in Lincoln, NE.