

Technical Advisory Committee – Tuesday, October 19, 2010
Embassy Suites Hotel
8:30 AM – 3:30 PM – Summary Notes

Present:

TAC members: Wayne Camara, Brian Gong, Linda Poole, Richard Sawyer, Dallas Watkins

NDE staff: Ed Foy, Ted Larson, Freida Lange, Jan Hoegh, John Moon, Marilyn Peterson (PM),
Pat Roschewski, Carol McClain (AM)

Governor's Office: Cheryll Wolff (AM), Matt Esch (PM)

State Board of Education: Bob Evnen (PM)

Contractor staff: Ron Mead , Patricia Johnson, Jessica Tickle(DRC), Dennis Ortman, Katie
Schoenhofer (CAL), Bill Auty (consultant to NDE)

8:30 AM Welcome and Introductions – Pat
Open Meetings Act Noted

8:45 AM Approve Summary Minutes from June 25, 2010 (Document 01) – Brian

- The Summary Minutes were approved.

8:55-9:30 Update: NeSA Reading Test Score Release (Document 02) – Pat – Dallas- Linda

- *The group discussed both the NeSA-R test results and the timing of the release. The timing delays were related to federal accountability requirements; the release of NeSA scores in August met the agreed-upon time line.*
- *In some districts the new test results were lower than under the old system; in other districts, the state test results mirrored the local test results.*
- *It was suggested that the raw score conversion charts for 2011 could be posted sooner so that districts testing the reading test online would have results earlier.*
- *The mathematics test release in 2011 will follow the same time line as the reading . Mathematics results should be released in August 2011, followed by a fall release of the accountability decisions.*
- *The TAC suggested that the NDE provide a tool for use in districts, especially those with small numbers. A chart could be developed to answer the question: "What is meaningful change?" This would help school leaders explain when differences in annual scores likely reflect real changes, and when those differences should be interpreted with*

much caution because there is high probability the differences may reflect fluctuations by chance.

- *The differences in performance between disaggregated groups are significant and will be the focus of school districts; a lot of attention is likely to be paid to these gaps.*
- *The group talked about the trends in serving students with disabilities and the group of students with autistic disabilities that is growing across the country.*
- *The TAC indicated the NDE should be thinking about how to help districts respond to the data. The NDE described the tools and workshops being prepared on “Leadership in the Use of NeSA Data.” These workshops are being delivered collaboratively with the regional service units and are scheduled to begin in November.*

9:30-10:00

Update: Long Range Assessment Plan and Item bank reports. (Document 03) - Jessica and Richard

- *After asking numerous questions about the long range assessment plan and the item bank report, the TAC suggested that the documents and the charts outlining the long range assessment plan and the item bank reports could be improved with clarification. The TAC recommended that the current documentation be expanded to reflect the interaction between matrix field-test forms, item yield, passage-dependent items, equating forms design, and the number of items needed over time to match the test blueprints.*
- *The TAC asked DRC to develop a narrative and a flow chart to accompany the expanded charts. These explanatory additions should describe the processes used in the long range planning for item development, forms construction, and the equating processes. The documentation should clearly allow NDE to track very specifically against the test blueprints what has been developed and what will be developed. For example, the documentation should distinguish between items that “are in the pool” and those that have completed field-testing satisfactorily and have met the published psychometric item quality guidelines. The planning documentation should clearly show that the targeted number of items (and not significantly more) are on track to be developed, according to the test blueprints.*

- *The TAC did not believe it was necessary that these accompanying narratives be developed for the 2010 technical manual, but wanted to be sure that NDE had all of the necessary information for the NCLB peer review, for NDE's own records (for example, if NDE had a question, to enable close contract review, or for the future if there were a change in contractors), and for future technical manuals. The TAC recommended that a general version be included in the technical manual, with more detailed versions being kept secure to help minimize the threat of "gaming" the system.*
- *The TAC suggested that NDE decide whether it wanted to cease developing items after the targeted number of forms were developed (2013 according to the presented plan), or whether it desired to develop additional items and possibly forms.*

10:00-10:15 Break

10:15-11:30 Update: The NeSA-R Technical Report (Document 04 – zip drive) – Richard, Wayne, DRC
 The Final Standard Setting Report (Document 05 – zip drive) – Wayne and DRC
 The NeSA Mode Comparability Report (Document 06) – Brian and CAL.

Assorted Items in the Technical Manual

- *The TAC suggested that for future clarity, the technical manual should explain that although the high school standards are written for "up to grade 12," the high school testing occurs in grade 11. For this year's peer review, this explanation will need to be covered in the overall system description.*
- *The TAC supported DRC's use of both pre-equating and post-equating checks. The TAC recommended that these equating designs and procedures be documented more clearly starting with the 2011 manual. This recommendation arose from the discussion about a basic pre-equating design, but the inclusion of several items for "core equating" form-to-form (year to year). The documentation should also include the equating design for computer-administered tests, since the online test forms differ from the paper-and-pencil forms. Some TAC members were interested in the experience of DRC and CAL of position effects on equating, since some studies have found equating items sensitive to small changes (e.g., edits and changes in position in the test), but the online test form design randomizes reading passage order (and items within passage?). Equating design and equating results have a bearing on the comparability studies discussed below.*

- *The TAC suggested that in the future technical manuals, narrative be included to explain how items with DIF were handled—this should include both a more complete description of the process of committee review after items are statistically identified, and it should include a summary of the actual results of the DIF review for the assessments being documented in the technical manual (e.g., the 2011 technical manual would include test development DIF reviews conducted in 2010-11).*

Standard Setting Documentation

- *It was suggested that in future, it would be helpful to provide more specificity in “who the panelists” were in the standard setting, i.e. exactly in what grade levels were panelists experienced. Wayne indicated he would provide a summary document indicating the changes that should be made to future technical manuals.*
- *The TAC recommended that the final standard setting documentation in the future include information about how the final standards came to be, including a description of the processes followed by the state Board of Education, information they considered, and so on, at a level of detail comparable to the other standard setting information, if possible.*
- *The TAC supported a survey of panelists following Standard Setting, as found in Section 5.4 in the 2010 Technical Manual, “Panelists’ Survey Evaluation Results.” The TAC recommended that future standard setting reports expand the documentation to include the results of the surveys as well as some interpretive text.*
- *Overall the standard setting process went very smoothly and Pat indicated the importance of the February session of stakeholders explaining the process ahead of time. That should be repeated for mathematics in February of 2011 if possible.*
- *The TAC indicated that the standard setting process that had been thoroughly discussed at many of the TAC meetings was successful, and the report was well done. Repeating the same standard setting processes would be a strong plan in 2011 for mathematics.*

Comparability Study

- *Pat explained that the comparability study was just a beginning step, that the 2011 comparability study for mathematics would be much more comprehensive, and that such a plan would be developed.*

- *The TAC acknowledged that this was an “initial study” but somewhat “superficial” and the results have “little value ” due to limitations in the study’s design that reflected only being able to analyze existent data. The group had numerous suggestions for improvement in 2011 including the following:*
 - *The comparability study should focus on comparability of the total scores when administration mode is considered. To do this, the study must be designed to detect total score differences. The contrast groups must be matched on significant variables, preferably an external achievement criterion. Work should be done ahead of time to design the comparability study and ensure that the needed data are collected. DIF studies at the item level are not sufficient to detect potential total score effects since most DIF methods assume comparability of performance on the total score by the two groups being compared for DIF. Mode interactions with significant subgroups should be examined as sample size and other resources permit. The purpose of these comparability studies is to provide NDE with documentation of the fairness of the program so it will be prepared prior to the program being challenged.*
- *The study should report on the practical effects of mode issues, such as impact on misclassification rates. The report language should be neutral, rather than advocating a particular position.*
- *However, to balance their recommendations regarding the design of a comparability study, the TAC suggested that NDE examine barriers to schools going online and to move quickly if the state is planning to go totally online. The TAC was generally supportive of moving to only computer-based testing as the state desires, so as to eliminate the need for comparability studies, to reduce costs of developing and administering multiple modes, and to be able to take full advantage of the strengths of computer-based testing.*
- *Some TAC members personally recommended that if the state had very limited resources, it should devote most resources to a rapid (i.e., in the next year or two) and successful transition to all computer-based testing, rather than devoting large resources to comparability studies. The TAC recognized that NDE might be subject to federal Peer Review requirements regarding documentation of comparability.*

11:30-12:00 Lunch – Begin the discussion of Statewide Writing

12:00-1:15 Transition of the NeSA-writing assessment – (Documents 07, 08, 09) Pat, Linda, and Dallas

- *The NDE shared a description of what has been occurring with statewide writing in the last 10 years, the reasons for proposed change and the transition steps being planned.*
- *The TAC advised careful revision of the scoring rubrics in the transition to analytic scoring so that everyone is sure that the criteria in the rubric are the criteria to be scored.*
- *The TAC advised that consideration be given to how the final score is to be derived, given the weighting of traits. The PLDs need to clearly articulate desired skills.*
- *The TAC advised careful consideration be given to standard setting methodology --- specifically, consideration should be given to a formula methodology based on the weighting of traits.*
- *The TAC discussed the value of third reads and discussed whether the best value of third reads is placed on domain sub scoring or total scores. The TAC's main concern was that NDE design the scoring model and the expectations in the field to balance appropriate scoring accuracy and financial resources.*
- *The TAC discussed the value of a second essay and artificial intelligence scoring, both of which are proposed in the Common Core Assessment in future years. Neither of these approaches are being considered in Nebraska at this time.*
- *Regarding the subject of testing 4th graders online, the TAC suggested conducting a survey to discover what other states are doing in regard to keyboarding below grade 8.*
- *In addition to the survey, the NDE and DRC want to have a conversation about a study that will provide some Nebraska information about the feasibility of testing writing with 4th graders on line, since many apocryphal stories are told about children keyboarding at young ages for non-testing situations such as texting, emailing, gaming, and social networking .*
- *Some TAC members asserted that "college ready" assumes being proficient in using a computer, including keyboarding. The state is assuming that students should largely have these competencies by grade 11, without formal instruction in school beyond what schools are currently providing (e.g., very short tutorials on specific features of the online testing platform, interface, etc.). If these competencies really are required to be "college- and career-ready," then at what point should students acquire them—Grade 11? Grade 8? Grade 4? The point is that if students need these competencies, the state should decide when to assume the students have these competencies, and whether these*

competencies need to be made part of the standards adopted by the state (currently the state assumes that students will acquire the competencies without a formal curriculum required by the state).

1:15-1:30 Break

1:30-3:00 State Accountability Model (Documents 10, 11,12) – Pat, Bill, and Brian

- *Bob, Bill, and Pat presented the work that the state board subcommittee had accomplished in the last several meetings. The framework document representing the work has been divided into three sections. Part One is the Beliefs, Values, and Purposes. The Second Part is Design and Implementation, and Part Three is the Time line.*
- *During the discussion additional points were offered as being important in the Beliefs, Values, and Purposes section. The accountability model should:*

Combine different sources of data.

Be based on the results or outputs of a school rather than the inputs or decisions made in that school.

Promote high achievement and shed light on actual performance.

Share the best practices of high performing schools across the state.

Emphasize growth in achievement and improvement over time.

Encourage high achieving schools and districts to improve while not penalizing schools that maintain high levels of achievement.

Use public reporting as a source of accountability and acknowledge that there are many important aspects of high quality schools that do not need to be included in accountability calculations.

Reduce opportunities for “gaming” the system.

Be easily understood and as “simple” as it can possibly be.

- *Brian suggested that the above points be added to Part One of the document and that the subcommittee consider including them because when it comes time to finalize implementation, the board will want to be specific about what it believes.*
- *Brian further suggested that it is appropriate for the whole board to adopt Part One at this time and wait on Part Two so that full implementation details can be verified with data.*
- *A discussion began on Part Two: Design and Implementation. Further work is needed to be sure that the narrative in Part Two matches the data scenarios that were provided. Bill Auty will be providing those revisions.*
- *The TAC pointed out the importance of agreeing on the performance indicators and their weight, so that the results of the computations reflect the values and beliefs, and so that data can be run to determine the impact of the decisions.*
- *The State Board of Education is intending to apply the first accountability decisions in 2011 with two years of reading data, growth in reading, the first year of baseline mathematics data, participation, and graduation data.*
- *The board accountability subcommittee work will proceed with Parts One of the Accountability Framework (with the additions from the TAC meeting) and with the Timeline, leaving Part Two – the design and the implementation details to be run, validated, and confirmed before fall.*

3:00-3:30 Wrap up and next steps –

Overall, the TAC indicated that considerable high quality work has been accomplished. The Nebraska Accountability System is of quality and meets the expectations outlined in the legislation.

Next meeting dates

Tuesday May 10, 2011

Tuesday October 18, 2011

