

Technical Advisory Committee  
Nebraska Department of Education  
December 4, 2013  
Cornhusker Marriott Hotel Lincoln, NE 8:30 am-3:00 pm  
**Minutes**

- 8:30 AM**      **Welcome and Introductions- Valorie**  
New Appointment to TAC  
Dr. Chad Buckendahl- Vita Included (Document 1)  
Reappointment to TAC  
Superintendent- Frank Harwood, Bellevue Public Schools
- 8:40 AM**      **Approve Minutes of April 24, 2013– Brian** (Document 2)
- 8:40-9:45**      **Update: Check for Learning- Jeremy** (Document 3)

Check 4 Learning [C4L] is in its third year of use with 176 districts that have signed the Memorandum of Understanding and use the system. NDE has also partnered with ESUs in order to support district efforts. The item bank will contain approximately 10,000 items by the end of the year. NDE provides this interim assessment tool for any Nebraska district but does not prescribe a specific type or frequency of use. Ultimately, NDE would like to determine effective ways in which Check4Learning system is being utilized by districts to improve student learning. NDE is partnering with Regional Education Laboratory Central [REL Central] to conduct a survey of districts and users of the Check4Learning system.

We are working together to determine:

- At what level C4L is managed in districts.
- The uses of the C4L system and the frequency of use.
- The uses of C4L data and the frequency of use.
- The structure of C4L professional development.

Project Timeline

- January 2014- Finalize surveys
- February 2014- Pilot the survey
- May 2014- Conduct surveys electronically
- June & July 2014- Analyze survey data

Check4Learning information and proposed surveys included as Document 3.

**Question:** *What input does the TAC have for content of the surveys? Are the data being gathered useful to identify possible districts with which to partner for further research?*

- Some additional questions were suggested for survey. If a district indicates it is not using Check for Learning, ask why that choice was made.
- Suggestion was made to include analysis of size of district using C4L and usage.
- Making C4L data available through the Data Dashboard would enhance both programs—and eliminates needless steps on school districts' part to make use of all data connected to NeSA scores.

- Other suggestions included: exploring how teachers are using the system to construct assessments, connecting C4L results to NeSA test score results, and including a school level survey.
- The district level survey could be adapted to be school level.
- NDE could encourage use of C4L as a measure of student learning objectives.
- Survey needs to focus on C4L as formative assessment.
- Some of the information asked on the survey could be collected as part of the set-up of the system that each district would complete for entry into the program.

**9:45-10:00      Break**

**10:00-11:30      Nebraska Accountability - Valorie (Document 4)**

#### Accountability

Nebraska Statute 79-760-06 established a legal requirement for an accountability model derived from NeSA scores, participation in NeSA and graduation rates. Based on the statute, the current NePAS system was adopted by the Nebraska State Board of Education in August 2012. In 2012-2013 Nebraska legislative session, LB 438 was introduced and is currently On File. If passed in 2014 in its current form, LB 438 would move accountability to a model that combines multiple measures to determine levels of performance for schools and districts. The direction of discussion and/or action of the Nebraska State Board of Education and Nebraska Legislators suggests that NDE research best practices of accountability models in order to be prepared to effect change. Indicators considered for inclusion by the Nebraska State School Board Ad Hoc Accountability Committee are included as Document 4.

**Question:** *Looking at the indicators globally, does TAC see the indicators as a sufficient and effective collection for school/district accountability? Does the TAC have suggestions for indicators not included or see any inherent difficulties in suggested indicators and/or their measure? What guidance can TAC share on use of Median Student Growth Percentile [SGP] or Median Student Adequate Growth Percentile [AGP] or a combination of both? Can TAC share technical advantages of basing status and growth measures on scale scores versus percent proficient? What are the advantages or disadvantages of setting goals within an accountability system?*

- Measuring growth may eliminate the need to focus on poverty as a subgroup. The classifications should be criterion-referenced. While weighting of improvement and growth can mitigate some of the unfairness in demographics, there are limits to how much it will do. Adjusting indicators for background characteristics leads to policy issues.
- Cut scores should be set and maintained over time.
- Even if the state accountability system is not a match to the Federal accountability system, resulting in two systems, using the same terminology and language in both may eliminate some confusion.
- Accountability is meaningless if districts are not looking at their own data and acting on it.

- Indicators are quite standard in any system that values college and career ready. Nebraska standards are being written to be college and career ready, using the ACT definition of readiness.
- Build a system—and plan to add indicators after the basic system for accountability is built. Could build in a system that changes the weighting of indicators as the system progress over years—with advance notice that the system will work like that.
- Size of school/ district will make a difference on whether districts are able to offer additional indicators, such as Advanced Placement, International Baccalaureate, and industry certification.
- Create definitions of a “good high school” --and explore the universality of the definition.
- Minimum group size affects the reliability of scores changes, but looking at the trend data will cancel out the year-to-year changes due to chance. Detecting small score change is hard to do reliably , but measuring change across multiple years helps stabilize the information.
- Confidence interval is more of a small school advantage and a disadvantage to larger schools. Should have larger and smaller school distinctions.
- Inclusion and precision are a balance—does NDE want to say the system is being set up to know schools can’t be more certain than the information allows? Information is more precise for larger schools.
- In terms of completeness of indicators, some states are looking at other content areas and the need to support other areas such as art. A good school does things in addition to RMSW, and these may be a distinction in classifying schools. School quality is everything
- Some states are looking at school climate.

**11:30-12:30 Working Lunch—Small Group Discussion (Document 5)**

Nebraska Department of Education is currently reviewing its accountability model-- NePAS. In addition, because Nebraska state statute requires a review of standards every five years, NDE Curriculum department has developed a plan for Nebraska Reading and Math Standards review. If Reading or Math standards are revised, NeSA tests will be reviewed for possible revision. We invite participants to hold discussions over lunch concerning their experiences with the current NeSA and NePAS systems through their work in the state and possibly in Nebraska school districts, the current atmosphere of testing and accountability in Nebraska and in the nation, and considerations to be included to provide effective leadership during times of transition. Proposed Schedule for Standards Review is included as Document 5.

**12:30-1:00 Input from TAC on transition**

**Question:** *What guidance can TAC offer to NDE that may contribute to smooth transition should the accountability model and/or state tests change?*

- If possible, states should avoid changing too frequently as it takes a lot of political maintenance to do so. A three-year span minimum is suggested for change to a new accountability system so schools and districts can grow into it.
- Stabilize the timeline. Make more changes at once, then wait longer for changes.
- Messaging and the way in which NDE characterizes the change are important. Many states use a public relations team to develop multiplatform communication platform. If a state department

shows what it is working towards, what changes are going to take place, and reasons for the change—transition is smoother.

- How NDE communicates is a critical piece. The department must be more diverse in how it communicates- engaging with key resources to help shape the message, so key partnerships are tapped and not just reacting to what NDE sends out and not just using traditional communication. Align with media representatives to expose the message.
- When change must take place in an accountability system, it is possible to use the same system, introduce a new parameter, and put out notice of weight shift.
- It helps to let people know where the system is going and allows districts to formulate resources. Getting framework set up and phasing in each part is part of stabilization.
- The bigger the change the more thoughtful about the transition NDE needs to be.

#### **1:00-2:45 Update: Nebraska Performance Accountability System [NePAS] - Valorie (Document 6)**

In November of 2010, the State Board of Education voted to approve the Framework of State Accountability that identifies the beliefs, values and purposes of such a system. In August of 2012 the Nebraska State Board of Education approved a policy to initiate an accountability system based on status, growth, and improvement, which are derived from NeSA scores, as well as participation, and graduation rate.

Four charts that include multiple indicators and rankings are provided for each school district, based on grade level configurations:

Chart 1: Grades 3-5

Chart 2: Grades 6-8

Chart 3: Grades 9-12

Chart 4: District

NePAS sample charts and Framework of State Accountability included as Document 6

#### **NePAS- A- Minimum Number- (Document 7)**

The decision was made to include in rankings any group with ten or more students. At some grade configurations in 2013, status scores for NeSA-Reading and NeSA-Math included 249 districts out of 249, which had enough students to be included in rankings. For NeSA-Science, which is administered in grades 5, 8, and 11, and in writing, which is administered in grades 4, 8, and 11, fewer schools are reportable for 2013 due to student numbers fewer than ten. Research on N is included as Document 7.

**NePAS-Question A:** *Given that a smaller N size is more variable, but includes more schools, what minimum N size would TAC recommend for inclusion on NePAS?*

- An N of 30 seems very high as about 25% of schools would not be included. NDE could lower the minimum N and use a confidence interval, but must consider where to draw the confidence level at.
- The higher the stakes the surer NDE will want to be about the information in its accountability system. One way to be surer is to get more information and take into consideration the certainty of information. In Massachusetts, there are three levels of consequences. The state used numerical for the first level and then schools submitted documentation about their improvement plans for the additional levels.
- Seven-year averages are needed to make a good decision. Need at least 3-5 year average to

make the first pass. For the second pass, look at more subjectively which schools will be the best investment. Fluctuations year to year won't be as impactful.

- Design the system for overall effect, then take care of special cases. Consider the three lowest-performing schools to be a special case. There is no way to set up a numerical system to identify three schools. Impose a second rule on top of the numerical system that makes sense.

#### **NePAS- B. NeSA Scale Score Growth- (Document 8)**

The Nebraska Board of Education has set simplicity as a priority for all elements of the accountability system. To that end, growth for individual students has been defined as the difference between this year's score and last year's score. The indicator for school accountability is the average of those differences for students enrolled in the school (matched sample). NeSA does not have a vertical scale; however the scale scores at every grade are set such that the range is 0 to 200 with 85 as the proficient cut score and 135 as the advanced cut score. Using the average of differences in scale scores across grades as a school accountability measure rests on the assumption that each scale score and difference score has a similar meaning across grades and specifically that a difference score of zero indicates roughly a year's growth.

Research on growth is included as Document 8.

- **NePAS- Question B:** *Given the Board's intentions and the NeSA data provided, are NDE's assumptions and use of difference scores as an indicator in NePAS reasonable? Does the TAC have suggestions for additional analyses?*
- More complex systems than simple growth give a better indicator of true growth
- In the current scale score system, there is not equivalence between scale score points.
- Investigate using Z scores.

#### **NePAS- C.- Measure of Growth for NeSA-Writing (Document 9)**

The 2012-2013 NeSA- Writing testing window took place January 21 through February 8, 2013. The test engine on the writing test caused some formatting errors for students in grades eight and eleven. NDE received reports of considerable frustration on the part of some students due to the formatting. The on-line issue was not resolved over the course of the three-week window. For the evaluation of the essays, Data Recognition Corporation [DRC] created a training packet and trained the DRC graders for a "fifth domain," in which the essay evaluators indicated for each essay if a formatting error appeared. The "fifth domain" did not contribute to the essay's score but merely serve as identification of essays that had formatting errors. Graders were directed to evaluate the essays without penalizing students for any formatting errors, and if any grader had a problem in evaluating a particular essay due to formatting errors to pass the essay to a team leader. "Presentation" is not a domain by which essays have ever been evaluated on NeSA-Writing.

Based on advice from the Technical Advisory Committee at the April 24<sup>th</sup>, 2013 meeting, NDE made the following decisions:

- To release the NeSA-Writing scores for Grades 8 and 11.
- To indicate on the score release reports and on the State of the Schools Report that formatting issues occurred in the on-line engine at grades 8 and 11.
- To **not** include the NeSA-W status scores for grades 8 and 11 in Nebraska Performance Accountability System [NePAS].

- To **not** include the NeSA-W scores in improvement for grades 8 and 11 in NePAS for the 2013 State of the Schools Report.

Statement included on score release reports and on the State of the Schools Report

^ Students at grades 8 and 11 experienced formatting issues with the NeSA-Writing online test administration. While research into the score results does not indicate an effect on student results, it also does not assure there was no effect. Scores should be interpreted with caution and are not included in Nebraska Performance Accountability System [NePAS].

Summation of NeSA Writing results at Grades 8 and 11, NeSA Writing 2013 –Formatting Analysis (presented at April 24, 2013 TAC meeting), and NeSA Writing 2013-Formatting Analysis Update are included as Document 9.

**NePAS-Question C:** *What measure would TAC advise NDE use for improvement at Grades 8 and 11 of NeSA-Writing for the 2014 release of NeSA-Writing scores and State of the Schools Report?*

- Scores from NeSA-W spring 2012 and from spring 2013?
- Scores from NeSA-W spring 2011 and scores from 2013?

*Or would TAC advise NDE to delay inclusion of improvement for NeSA-W at Grades 8 and 11 until the spring of 2015, using 2014 and 2015 scores?*

- Using 2013 scores for accountability would be problematic. NDE could use 2012 and leave a space in reports with asterisks and explanation to allude to degree of growth.
- Skipping a year is hard for people to interpret. Not sure about using 2012--although maybe some information is better than none. How do 2012/2014 compare? One solution is to do nothing. Second solution is 12/14. Third solution is to compare 12/14 with some guidance on how to interpret or what to expect in the future.
- 12/14 is easily misinterpreted.

**2:45-3:00**

Wrap up and next steps.

Dates for future TAC meetings

○