DRAFT
12/11/13
A. Porter

M E M O R A N D U M

TO: Michael Muenks, Director, DESE Assessment

FROM: Andrew Porter, Chair, DESE Assessment Technical Advisory Committee

SUBJECT: Advice to DESE based on meeting of December 9 and 10, 2013

The meeting of the DESE assessment Technical Advisory Committee took place at the Renaissance St Louis Airport Hotel on December 9 from 12:00pm-5:30 pm and on December 10 from 8:00am to 11:00pm. Members of the committee in attendance were Bertha Doar, Director of Assessment and Accountability, St. Louis Public Schools; Robert Linn, Professor Emeritus, University of Colorado; Ron Mertz, Consultant Emeritus, St. Louis Public Schools; Barbara Plake, Professor Emeritus, University of Nebraska; Andy Porter, Chair, University of Pennsylvania; Ed Roeber, Consultant, Michigan State University and University of Wisconsin; and Phoebe Winter, Consultant.

In attendance from the Department were Susan Newbold, Shaun Bates, and Michael Muenks. Persons in attendance from CTB were Anita Benson, Jessalyn Smith, Brenda West, and Jake Parizek. Persons in attendance from ARC were Jonathan Henry and Lisa Sireno. Persons in attendance from Questar were Lei Yu.

Missouri Assessment Program Update

Michael Muenks provided an update on the assessment activity in Missouri. He began by providing an overview of the alternate assessment program, the contractor for which is ARC. There are assessments in reading and math in grades 3-8 and high school, and for science in grades 5 and 8. The approach continues as in the past, with teachers providing samples of evidence of students' performance. Virtually all of the students assessed are judged to be proficient or better. The alternate assessment in reading and math will be replaced with a new Dynamic Learning Maps online adaptive system, but the ARC-led assessment will continue in science.

The MAP program continues with CTB as the new contractor. The approach is an augmented Terra Nova assessment. The MAP assessment suspended performance events for two years, but is now using them again. Missouri is a member of the Smarter Balanced Assessment Consortium; the Smarter Balanced assessments will replace the augmented Terra Nova assessments in 2014/15. The CTB-led science assessment is being moved to an online format. Teachers are reviewing the assessment and there will be a full field test in the spring of 2013.

At the high school level, Missouri uses end-of-course assessments. Algebra I, English II, and Biology are the NCLB accountability required end-of-course assessments. In addition, there are end-of-course assessments for English I, Algebra II, Geometry, American History, and Government. Riverside was the original contractor and did the development work. The program is now being administered through Questar but will transition to CTB starting in 2014.

Muenks reported that the Missouri federal waiver of NCLB accountability requirements is due to expire. A draft of the second waiver request is due in March and will be considered at the March TAC meeting.

Muenks reported that there are a number of professional organizations, including both teacher unions, which are expressing enthusiasm for a dramatic cutback in the amount of student achievement testing. There seems to be among these organizations some resistance to teacher evaluation, growth modeling, and Common Core standards. They would like to see testing done in just grades 3, 8, and the end-of-course exams, but with interim assessments in the NCLB grades.

Plans for a post-secondary readiness assessment continue with a target date to be operational in 2015. The purpose of the assessment is descriptive, but at the individual student level. There are no plans to use the assessment for accountability purposes. The initiative is running into some difficulties in how to best assess career readiness. **The**

TAC recommended that the National Assessment Governing Board website be used as a source of information about feasibility and approaches to assessing career readiness. At one point, NAGB decided that assessing career readiness was not feasible, but apparently, they are pursuing this possibility once again.

End-of-Course Assessments Technical Report for 2012/13

Copies of the technical report were circulated to members of the TAC in advance of the meeting. Lei Yu of Questar walked the members of the TAC through the technical manual, highlighting changes from previous years. Section 5.6.1 describes efforts to reduce/eliminate efforts to cheat. It was also pointed out that there was substantial growth in performance in Biology, probably due to the fact that Biology scores are now used for accountability purposes.

The TAC commented on the report, chapter by chapter:

- **In the Executive Summary, combine Tables E1, 2, and 3 into a single table with three columns and add the sample sizes.**
- **Explain what it means that the scores are no longer banked.**
- **Generally, in the Executive Summary, refer to the relevant chapter for each set of observations and make sure that the Executive Summary is consistent with what is stated in the relevant chapter. For example, on page 12 of Chapter 1, edit to clarify the role of districts and the state board.**
- **In chapter 2, Table 2.9, delete the "reporting categories" below the table.**
- **On page 68, clarify the statistical criteria used in the operational forms construction process.**
- **Page 110, explain what happens with DIF flagged items. Are any dropped?**
- **Page 113, section 5.1, 3rd paragraph, last sentence, the first word should be Questar, not Question.**
- **Page 117, the top line of text should refer to quality assurance forms, not EOC forms.**
- **Page 129, clarify that DESE follows up as appropriate when material is missing.**

- **Page 132, section 6.3.3 should be rewritten to accurately describe what actually was done.**
- **In section 6.5.2.4 on validity, describe what was found concerning drift.**
- **In chapter 9, since the data are cross-sectional, don't describe change as "growth," call it "change."**
- **In Table 10.26, add the number of score points possible for each item to provide context for the very high percentages of agreement.**
- **Explain what the multiple rater data are.**
- **Add quadratic weighted Kappas.**
- **Add text to describe what is in Tables 10.27 and 10.28.**
- **Page 313, in the list of purposes and uses, when planning for next year's report consider dropping evaluating programs and adding a summary of what was learned about identifying students' strengths and weaknesses.**
- **Add a short final chapter that pulls together what was learned about reliability and validity and more generally the quality of the data produced by the EOC assessments.**

The TAC complimented Questar on an excellent technical report.

The TAC asked to consider at a future meeting: (a) how forms are developed for EOC assessments and (b) standards for future work in regards to automated scoring.

<u>MAP-A Technical Report for 2013</u>

Jon Henry of ARC walked members of the TAC through the technical manual, which had been distributed in advance of the meeting without appendices. Henry reported that the results and format are pretty much the same as in previous years.

The TAC commented that using the read behind expert reviewer data done on 35% of the students assessed should probably not be used to change students' scores,

since that results in an inconsistency between those students who had a read behind and those students who did not.

On page 46, there are two positive and three negative findings in the list. Some members of the TAC felt that the report should comment on the conclusions to reach from that list of positives and negatives.

The current MAP-A in reading and mathematics will be replaced next year with a Dynamic Learning Maps approach. The science assessment will continue as is. The state is preparing for the likely result that the Dynamic Learning Maps assessments will not yield as positive overall results for student performance as the current MAP-A assessment has. There is some concern among teachers that this will reflect negatively upon them and this is being considered by the Accountability Systems TAC.

Grade Level Technical Manual

A copy of the grade level assessments technical report for 2013 was circulated to members of the TAC in advance (Michael, I can't remember the name of the person who walked the TAC through the discussion of this manual – could you insert it?). Discussion began with highlighting changes from the previous technical reports.

Overall, the TAC congratulated CTB on an excellent report. The TAC did suggest the following:

- **In chapter 7 on test results, there should be some commitment to expressing in the reporting of students' results the uncertainty due to errors of measurement.**
- **In appendices B and C, the screenshots could be cleaned up by being cropped.**
- **Tables 9.11-13, the larger than 1.0 correlations are the result of having corrected for attenuation. Perhaps that correction should be more prominent in the text.**

Ed Roeber absolutely loved chapter 2.

Data Forensic Analyses

Jessalyn Smith of CTB walked members of the TAC through Powerpoint slides presenting CTB's suggestions for data forensics. Three approaches were presented: erasure analyses, similarity analyses, and gain analyses.

The TAC complimented CTB on an excellent presentation of options. The TAC recommends that DESE pursue the erasure analyses, DESE do its own gain analyses and that these two procedures be used in tandem. The TAC recommends that the erasure analyses be done on last year's data and presented at the March TAC meeting. At that point, the TAC will consider with DESE and the contractor issues of how results should be communicated to districts, schools, and teachers, and what standards should be used for flagging (e.g. three standard deviations or four). Prior to that meeting, if possible, CTB will circulate to members of the TAC a copy of the paper they have written on the similarities and differences of the results across different approaches to detecting possible cheating.

New Science EOC

DESE continues to be in conversations with the University of Iowa and now also with CTB about the development of a new end-of-course assessment in science. This assessment is to be operational for the high school graduation class of 2018 and has an estimated $3 million development price tag. Current thinking is that the focus of the assessment will be on science studied in the freshman and sophomore years: Earth Science, Physical Science, and Algebra Physics (i.e. Physics First). These are three year-long courses and there will be an end-of-course assessment for each of them.

Work is proceeding slowly. This year's budget includes only $120,000 for the effort. The assessment when operational will not be a graduation requirement.

2014/15 Missouri Assessment Program Updates

Michael Muenks of DESE walked the TAC through a Powerpoint presentation. He began by announcing that CTB was the contractor selected for Missouri assessment work in 2014/15. There will be a great deal of activity between now and then. In spring 2014, there will be online field tests and grade level assessments for English Language Arts and Mathematics for grades 3-11 and Science in grades 5 and 10. The field tests will select schools to participate for one grade level content area combination. At the elementary level, this will be the first experience in Missouri with online assessments.

There will also be field testing for MAP-A using the Dynamic Learning Maps approach for grades 3-8 and 11.

At the high school, Missouri will have the end-of-course assessments gravitate toward Smarter Balanced assessments.

Muenks announced that the per-student costs of assessment would increase by three-fold. The development work is done by the Smarter Balanced Assessment Consortium and the costs are $9.55 per student. The administration, scoring, and reporting would be done by CTB at a cost of approximately $30 per student for a total combined cost of $40 per student.

The turnaround time for end-of-course assessments is five business days. The turnaround for fifth and eighth grade science assessments will be ten business days. In both cases, performance and constructed response tasks are human scored.

The Smarter Balanced developed English Language Arts and Mathematics grade level and end of high school assessments include interim assessments. They are online and they are adaptive. The field test of these assessments will be important because many questions remain about the feasibility and quality of resulting information about student performance. One particularly troublesome aspect is the classroom activity that teachers are to conduct prior to assessment. So far, in other contexts, that has not worked well.

For MAP-A, as stated earlier, science will continue as the current status collection of evidence. In English, Language Arts, and Mathematics, the Dynamic Learning Maps approach will be introduced. In a pilot, 35% of students assessed needed to have the online responses entry done for them. Muenks reported that there is considerable trepidation on the part of teachers at the prospect of comprehensive change due technology, a different assessment and greater item difficulty. As teachers experience the assessment, however, they become less apprehensive.

At the end of this discussion, Muenks clarified that DESE has asked for $18.5 million additional dollars for assessment to make the 2014/15 assessment program possible. If those funds are not provided, he is uncertain as to what approach to assessment will be taken.

Closing

The next meeting of the TAC is scheduled for March 6 and 7, 2014 and the meeting after that, August 21 and 22, 2014. At the meeting, TAC meetings were scheduled for December 4 and 5 of 2015, March 23 and 24, 2015, and August 20 and 21, 2015.