

6/7/2017

To: Lisa Sireno and Shaun Bates

From: Andy Porter, Chair, Missouri Technical Advisory Committee

Subject: Minutes of Missouri TAC Meeting on June 2, 2017

The Missouri Technical Advisory Committee met by telephone conference call on June 2, 2017. Members of the TAC in attendance were Bertha Doar, director of assessment, St. Louis Missouri Public Schools; Karla Egan, independent consultant; Ron Mertz, St. Louis public schools, retired; Barbara Plake, University of Nebraska, retired; Andy Porter, chair of TAC, University of Pennsylvania; Ed Roeber, independent consultant; and Phoebe Winter, independent consultant; DESE: Lisa Sireno, Shaun Bates Debbie Jameson and Lisa Scroggs; Questar: Elliot Dunn, Adam Johnson, Katie McLarty, Scott Bishop and Josh Borton.

Discussions focused on three issues from the spring 2017 administration of end of course tests: English I and II, algebra I and II, biology, geometry, government, American history, and physical science. Performance on English II, algebra I or II, biology and government are used for accountability purposes. Issue one concerned end of course testing on April 25, when heavy use of the system resulted in bandwidth limitations. Issue two concerned a software malfunction for the algebra I performance event. Issue three concerned the occurrence of students being timed out while responding to the English I and II writing tasks.

1. Impact of bandwidth limitations on April 25

Across all subjects, 14,666 students were recognized as either logged on and taking an assessment but with slow response time in moving from one item to the next or attempting to log on and being unsuccessful. At the time of the TAC meeting, the numbers in each of these two groups were unknown, but the number of students logged on and experiencing delays in moving from one item to the other is believed to be much the smaller group. Additional students of unknown numbers were unable to log on at all and therefore took the test at another time. In the text that follows these will be referred to as groups 1,2 and 3. Questar reported that the bandwidth problem was fixed within 90 minutes and that no student responses were lost. Testing windows were extended for approximately two dozen districts that requested an extension.

The question presented to the TAC was what course of action should be taken by DESE for accountability purposes. The discussion determined that student accountability was not a problem, since student scores on these tests are not used for student accountability. They can be used by teachers as one piece of information to determine a student's course grade, but this

is a teacher option and teachers did know which students were affected by the “slow down”. Thus, discussion focused on purposes of district and school accountability.

The TAC identified four options:

1. Use the obtained results as is for accountability purposes;
2. Do not use the results for accountability purposes;
3. Use some statistical procedure to adjust the results and use those adjusted results for accountability purposes;
4. Remove the affected students for one set of results for districts and schools and use the obtained results as another set of results; for each district and school, for accountability purposes use the better of their two results

No members of the TAC were in favor of statistical adjustment. Not using the results for accountability was not an attractive option, especially since EOC results have recently not been available for accountability. **Ultimately the TAC favored using for accountability option 4, the better results from the data as obtained and the data after affected students were removed. There was considerable discussion as to whether to remove only those students from group 1, those logged on and experiencing a serious slowdown in moving from one item to the next, or to also remove those from group 2, those that experienced the frustration of trying to log on and not being able to.** Unfortunately, the membership of group 3, those who couldn’t even register an attempt to log on is unknown. The TAC recognized that the frustration experienced by students in each of these three groups could have bearing on their performance on the affected test for them. Questar estimates that approximately half of the students in the state were possibly affected.

Though Questar presented contrasts in results for the various effects in groups from previous years and the current year, it was impossible to interpret these comparisons because the representativeness of the affected groups to students statewide was unknown. Questar estimated that they would have the analysis file necessary to make these representativeness determinations on or before June 19. **Some members of the TAC believed that only group 1 members might be used in the preferred option for accountability purposes if the students in group 2 were representative to statewide students. Other members of the TAC felt that both groups 1 and 2 should be used in the preferred option for accountability purposes regardless. The TAC will meet again virtually once Questar has produced the requested analyses to determine the representativeness of students in the affected groups. Not only did the TAC recommend that Questar produce results on representativeness in terms of prior-year test scores and demographics but also results on the percent of students in each affected group in aggregate and the distribution of percent of affected students in each group for schools. Questar was charged with providing an analysis plan to DESE by close of business the following Monday (the TAC meeting was on Friday).**

2. Software malfunction for the algebra I performance event

For 172 students, the performance event software for algebra I created unscorable responses. These unscorable responses were triggered by students entering a select few characters that for unknown reasons created gibberish for their entire answer. While this problem has been fixed for future administrations, the question to the TAC was what to do about the affected students. Questar presented the TAC in advance of the meeting two short papers addressing this issue. The Questar proposal was to statistically link student performance from session one which was not affected to student performance for the full test for those not affected. Then, to use the statistical link to produce the predicted results for the affected students. One of the handouts showed that the affected students tended to be very low performing students. The plan from Questar was to report student performance for both session 1 (having no performance event) and session 2 (the performance event worth 10 points).

The TAC concluded that the statistically estimated scores for the affected students in session 2 was appropriate and justified, conditioned on the correlation between predicted scores and actual scores for the unaffected students being high. DESE requested that Questar produce this correlation and in addition a distribution of the absolute value of the deviations of predicted and actual scores. The TAC will review these results and make a recommendation as to whether to go ahead with using the statistical predicted scores.

3. Students being timed out for the English I and II writing tasks

For some reason, students who spent a long time entering text in response to the writing tasks on the English I and II end of course assessments were timed out of the testing session. The students could sign back in to complete the test, but still they experienced frustration. Unfortunately, the number of students and who the students are who were affected is unknown. This problem occurred across platforms and is seemingly widespread yet known only anecdotally.

The TAC recommended that in the future students be given a warning when they are about to be kicked off of the system. Questar thought this may solve the problem for the future. As for the recent administration, the TAC had no solution to offer as the affected students are not known.