



Assistant Commissioner
Office of State Assessment

August 2019

TO: District Superintendents
Superintendents of Schools
Principals of Public, Religious, and Independent Schools
Leaders of Charter Schools

FROM: Steven E. Katz *Steven E. Katz*

SUBJECT: Comparability of Spring 2019 Grades 3-8 English Language Arts and Mathematics Paper-based and Computer-based Tests

The purpose of this memorandum is to provide information about the results of the comparability study that was conducted for the Spring 2019 Grades 3-8 English Language Arts (ELA) and Mathematics paper-based and computer-based tests.

Background

In Spring 2019, the Department offered the Grades 3-8 ELA and Mathematics Tests in two administration modes: paper-based testing (PBT) and computer-based testing (CBT). Administering these tests via CBT was optional for schools and those schools that chose to offer CBT made this decision independently for each subject and grade. The Department provided readiness verification tools to help those schools selecting CBT ensure they were well equipped and prepared to provide a successful CBT experience for their students. Additionally, several CBT practice test sessions were made available to CBT schools to familiarize students and teachers with the CBT delivery system. Each of the CBT practice test sessions featured examples of all of the types of test questions included on the tests. This provided the opportunity for students to practice answering both multiple-choice and constructed-response questions on the computer devices they would be using for the actual test.

To further ensure fairness, the Department's contractor, Questar Assessment Inc., conducted a comparability study to identify whether or not there were any differences in student performance that could be attributed to the mode of test administration (i.e., PBT versus CBT). The comparability study methodology and results are summarized below. The findings of this study were used to ensure that students received a score that was representative of their knowledge and skills, regardless of whether they took the tests on paper or computer.

Comparability Study Methodology

Only some schools chose to administer the tests via CBT (representing approximately 19% of all ELA test takers and 14% of all Math test takers). Therefore, the population of students who tested via CBT were not assumed equivalent to the population of students who tested via PBT. In order to select a sample of students who tested via PBT that could be compared to those students who tested via CBT, a method called propensity score matching was employed. Propensity score matching allowed for the identification of groups of students who tested via

PBT that were similar to the groups of students who tested via CBT on a number of school and student characteristics, including achievement on the prior year's test.

Using these characteristics, Questar selected a group of PBT students that matched the group of CBT students for each grade and subject. This allowed for a direct comparison of student results between the two groups. For comparison, the mean scale scores were calculated for each grade and subject by mode of testing. The results are shown in the section below.

Results of Comparability Study

Table 1 shows the scale score means for the PBT and CBT groups on the 2019 English Language Arts Tests by grade as well as the differences in mean scale scores between the matched groups. Table 2 shows these same data for the 2019 Mathematics Tests.

Table 1. *PBT and CBT Means and Differences for Grades 3-8 ELA*

	PBT Scale Score Mean	CBT Scale Score Mean	Difference (Rounded to nearest whole number)
Grade 3	See footnote*		n/a
Grade 4	598.2	595.7	+2
Grade 5	600.0	597.8	+2
Grade 6	598.7	597.4	+1
Grade 7	598.9	598.2	+1
Grade 8	598.9	598.0	+1

* Because Grade 3 students have no prior test results on which to match PBT to CBT students, a PBT comparison group was not created and group means were not calculated for this grade level.

Table 2. *PBT and CBT Means and Differences for Grades 3-8 Math*

	PBT Scale Score Mean	CBT Scale Score Mean	Difference (Rounded to nearest whole number)
Grade 3	See footnote*		n/a
Grade 4	599.7	598.4	+1
Grade 5	600.7	599.0	+2
Grade 6	601.7	600.1	+2
Grade 7	602.6	600.4	+2
Grade 8	599.3	597.7	+2

* Because Grade 3 students have no prior test results on which to match PBT to CBT students, a PBT comparison group was not created and group means were not calculated for this grade level.

For those tests in which a difference in mean scale scores between the two comparable groups was observed, the scale scores for all students who took the test in that grade via CBT, (which was the lower scoring mode in all instances during this administration), were adjusted by adding the number of scale score points shown in the “Difference” columns of Tables 1 and 2 to the CBT students’ scale scores, up to the maximum attainable scale score. Thus, the scale score adjustments for students who tested via CBT, shown in Table 3 below, reflect the differences between the PBT and CBT groups found in the comparability study. These slight adjustments ensured that students who demonstrated comparable proficiencies in their knowledge and skills received comparable scores whether they tested on paper or on computer.

Table 3. *Summary of Scale Score Adjustments for CBT*

	ELA Scale Score Adjustment	Math Scale Score Adjustment
Grade 3	+2*	+2*
Grade 4	+2	+1
Grade 5	+2	+2
Grade 6	+1	+2
Grade 7	+1	+2
Grade 8	+1	+2

* Because Grade 3 students have no prior test results on which to match PBT to CBT students, a PBT comparison group was not created and group means were not calculated for this grade level. Instead, the mean adjustment for the other elementary grades for which a comparison was possible (i.e., Grades 4 & 5) was applied to the scores of Grade 3 students who tested via CBT.

For questions concerning the Grades 3-8 ELA or Mathematics Tests, please email the [Office of State Assessment](#) at call 518-474-5902. For questions concerning CBT, please email [CBT Support](#).