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 was similar to the groups of students who tested via CBT on a number of school and student characteristics, including achievement .

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 2017 English Language Arts Tests by grade as well as the differences in mean scale scores between the

| | | | |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

| | | | |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |