

**SUMMARY OF REPORT OF INDEPENDENT INVESTIGATION
INTO ALLEGATIONS OF TAMPERING WITH SUMMER, 2008
CRCT MATHEMATICS RETEST AT DEERWOOD ACADEMY**

I. Background

The Atlanta Public Schools [APS], through the Office of the General Counsel and the Office of Internal Resolution [OIR], communicated with Penn Payne, LLC on June 19, 2009, and retained the firm to conduct an independent and unbiased investigation into allegations contained in the Preliminary Comprehensive Audit Report: Summer 2008 CRCT Retest Grade 5 ("the Report"), dated June, 2009 and prepared by the Governor's Office of Student Achievement ("GOSA"). The Report alleged that "someone who had access to test materials after testing concluded changed multiple students' answers on the 5th grade mathematics CRCT Retest ("the Retest") at Deerwood Academy ("Deerwood") in July, 2008. The Report alleged that the motivation for altering test answers was to enable enough Deerwood students to pass the test so as to achieve Adequate Yearly Progress ("AYP"). It stated: "Schools that do not make AYP ... suffer consequences," but without specifying those consequences. The Summer 2008 Retest was the first time that a Retest was allowed to be used for purposes of AYP.

Five APS elementary schools sent their students to the Deerwood Summer school: Perkerson, Fickett, Continental Colony, Kimberly, and Deerwood. Thirty-two Deerwood 5th grade students attended Summer school and took the Retest. All of them were students who had not passed the Spring CRCT testing in mathematics. The Report identified eleven of those 32 students through erasure analysis who were in the 98th-99th percentiles on total erasures and/or wrong-to-right changes of answers (the "flagged students"). The Report listed three indicators of test tampering: "[1] the results of ... erasure analysis, [2] the unusually high gains made over the Spring CRCT administration, and [3] the fact that the scores of the other four Summer school programs at Deerwood did not spike. GOSA finds overwhelming evidence to show that someone changed multiple students' answers." GOSA subsequently added a fourth indicator of tampering to those three—out of the Deerwood students who failed the Spring 2008 CRCT in 5th grade math and who retook the test in Summer school, nearly 50% failed the 6th grade math CRCT in the Spring of 2009.

The scope of my investigation was to determine whether there was evidence to support or contradict GOSA's allegations of tampering with the 5th grade mathematics Retest.

II. Certain Irregularities in the Testing Process Were the Result of Negligence and Were Not Deliberate Attempts to Alter Answers.

The investigation showed that there were some irregularities in the 2008 Summer Retest process at Deerwood Academy. Some of the procedures used there by the testing coordinator, Tabeeeka Jordan, and the site administrator, Lavonia Ferrell, were in

violation of the Georgia Department of Education instructions for administering the Retest. The flaws in the process of administering the tests gave several people the opportunity to alter answers on a small number of tests, although not necessarily the answers of the 11 students flagged by GOSA. I have concluded that the irregularities were the result of negligence in record keeping, lack of energy and diligence in following the rules, and supervisory failure, rather than being deliberate attempts to find opportunities to alter the test answers.

III. GOSA's Data Did Not Provide "Overwhelming Evidence" of Tampering.

GOSA did not attempt to determine how the Retest answers were altered or who had made the erasures and alterations, but simply determined that multiple students' answers had been changed, using the four indicators listed above as proof. Of those four indicators, I conclude as follows:

(1) The gains in Deerwood's scale scores can be explained by factors other than test tampering. Deerwood was a Teach for America site during the Summer school, and all Deerwood math students except one were in classes staffed with four college graduate teacher trainees, in addition to an experienced teacher. No math class had more than 15 students, resulting in an unusually favorable teacher to student ratio. The students received an extraordinary amount of individual attention in Summer school, far more than during the regular school year. Twenty-five of the 32 Deerwood students were deliberately placed in classes with the teachers whom the site administrator, Ms. Jordan, considered the best teachers, and she monitored those classes the first week of school. And for the first two weeks of Summer school, the Deerwood students were given extra homework packets on Friday afternoons.

(2) GOSA has declined to give me access to the original answer sheets of the Deerwood students, and without seeing them, I cannot evaluate GOSA's evidence based on erasure analysis. For example, I do not know if any of the answer sheets show a pattern of erasures to indicate that a student skipped a bubble for an answer by mistake, put succeeding answers in the bubble that was one ahead of where it should have been, had to go back and erase all the preceding answers up to the point of the skipped answer, and then put in the new answers where they belonged. I do not know if the opposite happened—with a student missing a question and putting answers one bubble behind where they should have been. Until APS personnel or I have access to the answer sheets, these questions, and other questions, create doubt about the reliability of the erasure analysis to indicate tampering.

(3) There is a partial explanation for the "spike" in Deerwood's math scores in contrast to the four other schools, because Ms. Jordan put the majority of Deerwood students with the best teachers and gave them some extra homework. There may also have been an error in GOSA's computation of Deerwood's gains from the Spring to the Summer Retest, if GOSA counted a decrease of 31 scale score points in one student's score as an increase. Nevertheless, this spike is the strongest and most persuasive statistical indicator of tampering.

(4) The comparison between some of the Deerwood students' 5th grade Retest scores and their 6th grade scores in the Spring of 2009 is invalid, based on the Georgia Department of Education's own CRCT Score Interpretation Guide, which states that "it is not appropriate to compare scale scores across grades and content areas." This indicator is therefore not at all persuasive as proof of tampering.

GOSA's four indicators provide some statistical evidence of tampering, but not "overwhelming" evidence. And when I take these indicators and combine them with the lack of evidence of how tampering was accomplished and who did it, the persuasiveness of the statistical indicators is diminished.

IV. There Was No Concrete Evidence That the People Who Had Access to the Test Materials Used Their Access to Change Students' Answers.

The test materials were stored in a conference room that connected to the principal's office through a door and that had another door to the hallway in the school's office suite. The investigation showed that Ms. Ferrell, the test coordinator, Ms. Lisa Smith, the school principal, Ms. Lashanna Morrison, the school administrative assistant, and Ms. Margaret Merckerson, a volunteer retired APS educator, had access to the completed answer sheets. Ms. Smith and Ms. Morrison had keys that allowed them access to the conference room, and Ms. Ferrell and Ms. Merckerson were given access during the testing period so that they could coordinate the testing and assist with the testing, respectively. Ms. Jordan, the site administrator, who was Deerwood's assistant principal, did not have keys to the conference room where the test materials were stored, or to the principal's office, but she should have been regarded as a person authorized to enter the room, according to the 2008 "School and System Test Coordinator's Manual," and she could therefore have gained access to the answer sheets.

Ms. Smith was not at Deerwood during the week of testing, except for the Monday of that week, and the test administration did not begin until Tuesday. She was not the principal on duty for that week and was not required to be there. Any scenario that places Ms. Smith at Deerwood after the tests were administered is speculation, and she could not have tampered with the tests without being physically present.

There is no concrete evidence that any of the other people who had access to the test materials used their access to change students' answers.

V. No One Who Had Access to the Tests Had a Sufficiently Strong Motivation to Achieve AYP to Take the Risk of Test Tampering.

Ms. Smith and Ms. Jordan were clearly motivated to make AYP, but not so strongly motivated that they would take the risk of tampering with the answer sheets. Making AYP in 2008 could have resulted in a small cash remuneration to the school, in an amount similar to the 2007 payment of \$824. But the direct financial benefit is far too small to be a motivating factor for anyone to tamper with the answer sheets.

Making AYP was a positive achievement for Ms. Smith's and Ms. Jordan's careers. For example, it was noted in Ms. Smith's 2007-2008 performance review that she had made AYP. But, since Deerwood had made AYP the year before the Retest, failure to make it in 2007-08 did not bring negative consequences from the State. Ms. Smith and Ms. Jordan still had the following year to make AYP so as to prevent Deerwood from being labeled a "Needs Improvement" school.

In judging Ms. Smith's and Ms. Jordan's motivation to tamper with test answers, it is important to distinguish making AYP from making a school's targets that are set by APS. The APS targets are calculated differently, using a different student population and test result thresholds, than AYP. Reaching at least 70% of the APS targets provides compensation to virtually all employees in a school, and Deerwood had done well in reaching its targets in previous years, although its scores had plummeted in 2008. But, significantly, only the regular school year administration of the test is used for reaching the APS targets, not the Summer school administration. Tampering with the tests in the 2008 Summer Retesting could not, therefore, change the fact that Deerwood had not come close to making its targets in 2007-2008. Nor could it change the fact that Ms. Smith was automatically put on a Performance Development Plan as a result of Deerwood's poor showing in reaching the APS targets.

The test coordinator and the two volunteer retired APS educators all served as substitute teachers at Deerwood in the 2008-2009 school year following the Retest, and that employment could theoretically have been an enticement for one of them to conspire with Ms. Smith or Ms. Jordan to alter the test results. But the evidence does not support a motive to conspire with others, because Ms. Smith did not need anyone else if she had wanted to tamper with the tests, and Ms. Jordan did not have the authority to hire the three educators. Moreover, none of these three women had an independent motivation for Deerwood to make AYP.

The administrative assistant had some minimal career motive to give the site administrator access to the room where the completed tests were stored, but there is no evidence that she did so. The teachers who taught the large majority of Deerwood students had no motivation whatsoever to help Deerwood make AYP, since they taught at other schools during the regular school year.

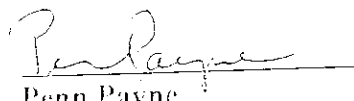
VI. There Was No Concrete Evidence of How the Alteration of Answer Sheets Would Have Been Accomplished.

It would not have been an easy task to change answers on the students' tests in a way that would result in an increased pass rate. A person tampering would need a test booklet and then would have to answer all the questions, mark the correct answers, possibly make their own answer key, identify the Deerwood students' tests from among the other four schools, decide which children's tests needed altering, decide how many questions to alter on each test, decide which questions to change by identifying wrong

answers on each test, physically make the erasures, and then bubble in the right answers. It would have taken several hours to do all of this activity. I did not find evidence that any of it was done.

VIII. Overall Conclusion

My overall conclusion is that, despite the irregularities in the testing process and despite some evidence of opportunities and motive to tamper with the tests, there is no concrete, non-statistical evidence that anyone actually altered answers or tampered with the answer sheets. Any contrary conclusion could only be based on speculation and on possibilities about what might have happened, not what can be proved to have happened. In a matter as serious as this one, I will not reach a conclusion of tampering without more persuasive evidence than I have found in this investigation.



Penn Payne

August 20, 2009