

**Statewide Analysis: Spring 2012 EOCT Courses***Presented to the State Board of Education on February 20, 2013***Rationale for this Investigation**

The Governor's Office of Student Achievement (GOSA) is charged with auditing and inspecting schools and Local Education Agencies (O.C.G.A. § 20-14-26). An analysis of the 2012 Spring EOCT answer documents conducted by the state's vendor, NCS Pearson, Inc., showed an unusually high number of answers changed from wrong to right (WTR) in some classrooms. Based on a conservative criterion for identifying unusual results, OSA makes the recommendations in this report to help eliminate test misconduct and to help students adversely affected where applicable.

Because important decisions for individual students and for schools are based on EOCT data, it is vital that scores provide an accurate representation of students' knowledge.

Purpose of the End-of-Course Test (EOCT)

The EOCT is a standardized assessment administered in Spring 2012 to students enrolled in high school core content courses. It is designed to measure how well students ascertain the knowledge and skills within the state's curriculum, the Georgia Performance Standards and provide diagnostic information to help students identify strengths and areas of need in learning.

Executive Summary**Erasure Analysis**

The state's testing vendor for the EOCT, NCS Pearson, Inc., conducted an erasure analysis on the Spring 2012 answer documents for the high school core content courses (Math I, Math II, GPS Algebra, GPS Geometry, US History, Economics, Biology, Physical Science, Ninth Grade Literature, and American Literature). The analysis examined student answer sheets for all EOCT test-takers in the state and identified classrooms where wrong answers were changed to right answers at above-average rates. The average classroom rate was determined by looking at the entire state testing population in each course.

Using a professional grade scanner, Pearson scanned the answer sheets to determine the total number of erasures and the total number of wrong-to-right (WTR) changes on each document. Pearson then aggregated those results at the classroom level. Any classroom in which the number of WTR changes was 3 standard deviations (SD) or more (adjusted for class size) above the state average for that particular course was "flagged" as having an unusually high number of WTR changes.

Erasure Analysis Results

Overall, 90% of students did not have an answer changed from wrong to right. Approximately, 96% of schools had less than 10% of their classrooms flagged or fewer than 5 of their classrooms flagged. While the majority of schools in the state have relatively low percentages of flagged classrooms, the analysis indicates that there are still a few classrooms across the state with unusually high number of wrong answers changed to right answers.

Rationale for this Report

The Governor's Office of Student Achievement (GOSA) serves as the reporting and accountability agency for education in Georgia. As such, it is charged by law (O.C.G.A § 20-14-26) with auditing and inspecting schools and Local Education Agencies. As the current vendor for the delivery of the End-of-Course Tests (EOCT), Pearson is providing services to GOSA to complete an Erasure Analysis for the Spring 2012 EOCT. This report is a summary of the comprehensive analysis completed on the Spring 2012 End of Course Tests (EOCT) main administration.

EOCT Assessment Overview

The A+ Educational Reform Act of 2000, O.C.G.A. §20-2-281, mandates that the State Board of Education adopt end-of-course assessments in grades nine through twelve for core subjects to be determined by the State Board of Education. The EOCT is administered upon the completion of Mathematics I, Mathematics II, GPS Algebra, GPS Geometry, United States History, Economics, Biology, Physical Science, Ninth Grade Literature, and American Literature (a total of 10 EOCT). It is designed to measure both the effectiveness of classroom instruction at the school, system, and state levels and the strengths and areas of need in learning for examinees. A student's final grade in the applicable course is calculated as follows (State Board Rule 160-4-2-.13):

- For students enrolled in grade nine for the first time before July 1, 2011, the EOCT counts as 15% of their final grade.
- For students enrolled in grade nine for the first time on, or after, July 1, 2011, the EOCT counts as 20% of their final grade.

Georgia school systems have the option to administer the test in one of three ways:

- Paper and pencil administration with answer documents sent to Pearson for scoring
- Paper and pencil administration with answer documents scanned locally with software developed by Pearson
- Online administration with electronic responses scored at Pearson

The EOCT is part of Georgia's high school accountability assessment and is used as part of the College and Career Readiness Performance Index (CCRPI). Important decisions for individual students and for schools are based on EOCT data. Therefore, it is critical that reported scores are an accurate representation of students' knowledge.

Erasure Analysis

Pearson conducted an erasure analysis for the Spring 2012 EOCT main administration for assessments completed via paper/pencil. The analysis was conducted for EOCT in Mathematics I, Mathematics II, GPS Algebra, GPS Geometry, United States History, Economics, Biology, Physical Science, Ninth Grade Literature and Composition, and American Literature and Composition. The purpose was to identify classrooms where item responses were changed more frequently when compared to the typical EOCT classroom for the state test population.

Erasure Capture Method

Pearson uses optical mark scanners (OMR) to capture data from the scannable forms used on the EOCT. Scanners have the ability to discern between pre-printed coding and respondent markings using a 16-level mark discrimination system. An erasure, for paper testing, is determined by the following criteria: The highest intensity mark on the answer document is automatically classified as the examinee's response while the discernible mark with second darkest intensity is classified as the erasure. All scanned EOCT answer documents were analyzed using the mark discrimination system to determine responses that changed from wrong to right (WTR) and total erasures (ERA) on each answer document.

Statistical Method

The method used to analyze the erasure data uses state and classroom population mean and variance adjusted for class size. The flagging procedure was applied to all erasures and wrong to right erasures. The erasure analysis utilized data from all test items, including embedded field test items.

The statistical test used for flagging is based on a test of the null hypothesis (H_0) that the mean number of erasures for a class is drawn from a random sample from the state distribution of erasures for a class. The alternative hypothesis (H_1) is that the mean number of erasures for a class is too high to be the result of a random sample. Classes that are flagged due to the rejection of H_0 should be further analyzed to see if there is a non-random explanation for the flag.

The central limit theorem holds that the sampling distribution of a mean number of erasures for class c (m_c) is asymptotically normal with mean and standard deviation

$$\text{Mean}(m_c) = \mu \quad (1)$$

$$\text{SD}(m_c) = \frac{\sigma}{\sqrt{n_c}} \quad (2)$$

with n_c denoting the size of class c and m_c denoting the mean number of erasures for class c . Additionally, μ and σ denote the mean and standard deviation of the number of erasures of the population of examinees taking the EOCT in Georgia.

Classes were flagged if the m_c was larger than $\mu + 3 \frac{\sigma}{\sqrt{n_c}}$. Dividing the standard deviation by the square-root of n (Equation 2) allows the statistical test to be sensitive to different class sizes. For example, if the state mean and standard deviation are 2.34 and 3.65 respectively, the flagging criterion for a class size of 16 would be $\left(2.34 + 3 \frac{3.65}{\sqrt{16}}\right) = 5.08$, while the flagging criterion for a class size of 36 would be $\left(2.34 + 3 \frac{3.65}{\sqrt{36}}\right) = 4.17$.

The flagging criterion was set at 3σ to minimize the probability of false positive errors (Type I) in the statistical analysis. Under a random sampling of a normally distributed variable, the standard normal table shows that the probability of a sample mean being more than three standard deviations above the population mean is approximately 0.001. Rejection of H_0 only shows that the observed mean number of erasures for that particular class is unlikely to be the result of random sampling.

Erasure Analysis Results

The Spring 2012 EOCT Erasure Analysis is the first official erasure analysis completed for the paper administration of the EOCT. Key initial findings from the erasure analysis include:

- The mean number of wrong to right erasures ranged from 0.490 to 0.966 across the 10 EOCT subjects. On average, approximately 0 to 1 wrong answers were erased and then correctly answered per examinee per answer sheet.
- 66.27% of schools were not flagged in any subjects.
- 68 of the 593 schools (11.5%) had at least one paper EOCT flagged for a higher than expected number of wrong to right erasures for at least 10% of their classrooms.
 - Of those 68 schools, 30 schools had only one classroom flagged, 7 schools had two classrooms flagged, 5 schools had three classrooms flagged, and 4 schools had four classrooms flagged.
- 22 of 593 schools (3.7%) had at least five classrooms flagged for wrong to right erasures and at least 10% of their classrooms were flagged for wrong to right erasures.
- Approximately 96% of the schools had either less than 10% of their classrooms flagged or fewer than five of their classrooms flagged for a higher than expected number of wrong to right erasures.

Spring 2012 EOCT Main Administration Erasure Analysis Summary Tables

State Summary Statistics for Total Erasures by EOCT Subject

Table 1 displays a summary of the erasure data for EOCT tests administered. The table includes subject area, total number of examinees (N), total number of erasures, the mean and standard deviation for total number of erasures, the correlation between the total number of erasures and wrong to right erasures, the number of erasures by percentile and the maximum number of erasures. The mean number of erasures ranged from 0.490 to 0.966 across the 10 EOCT subjects. Approximately 0 to 1 responses were erased per examinee per answer sheet. At the 90th percentile, the erasure count was between 2 and 3, which means that 90% of examinees had less than that number of erasures for the respective EOCT subject.

Table 1. State Summary Statistics for Total Erasures by EOCT Subject

| Subject Area | N | Number of Erasures | Mean | Standard Deviation | Correlation with WTR Erasures | Number of Erasures by Percentiles | | | | | | Maximum Number of Erasures |
|------------------|-------|--------------------|-------|--------------------|-------------------------------|-----------------------------------|----|----|----|----|------|----------------------------|
| | | | | | | 50 | 75 | 90 | 95 | 99 | 99.9 | |
| Mathematics I | 74846 | 36689 | 0.490 | 1.177 | 0.799 | 0 | 1 | 2 | 3 | 5 | 11 | 29 |
| Mathematics II | 70099 | 40945 | 0.584 | 1.357 | 0.805 | 0 | 1 | 2 | 3 | 6 | 13 | 30 |
| GPS Algebra | 18350 | 17733 | 0.966 | 1.693 | 0.817 | 0 | 1 | 3 | 4 | 8 | 14 | 24 |
| GPS Geometry | 7372 | 6700 | 0.909 | 1.768 | 0.844 | 0 | 1 | 3 | 4 | 8 | 15 | 25 |
| US History | 57705 | 40320 | 0.699 | 1.565 | 0.844 | 0 | 1 | 2 | 3 | 7 | 15 | 36 |
| Economics | 29370 | 26302 | 0.896 | 1.817 | 0.844 | 0 | 1 | 3 | 4 | 8 | 17 | 39 |
| Biology | 68259 | 51742 | 0.758 | 1.544 | 0.829 | 0 | 1 | 2 | 4 | 7 | 14 | 31 |
| Physical Science | 44909 | 41537 | 0.925 | 1.700 | 0.827 | 0 | 1 | 3 | 4 | 8 | 15 | 33 |
| 9th Grade Lit. | 76001 | 48260 | 0.635 | 1.370 | 0.848 | 0 | 1 | 2 | 3 | 6 | 12 | 34 |
| American Lit. | 59837 | 48157 | 0.805 | 1.564 | 0.852 | 0 | 1 | 3 | 4 | 7 | 14 | 35 |

State Summary Statistics for Total WTR Erasures by EOCT Subject

Table 2 displays a summary of the **wrong to right erasures** for EOCT tests administered. The table includes subject area, total number of examinees (N), total number of wrong to right erasures, the mean and standard deviation for the total number of wrong to right erasures, the correlation between the total number of erasures and wrong to right erasures, the number of wrong to right erasures by percentile, and the maximum number of wrong to right erasures. The mean number of wrong to right erasures ranged from 0.213 to 0.439 across the 10 EOCT subjects. Approximately 0 to 1 wrong responses were erased and then correctly answered per examinee per answer sheet. At the 90th percentile, the wrong to right erasure count was 1, which means that 90% of examinees had less than that number of wrong to right erasures for the respective EOCT subject.

Table 2. State Summary Statistics for Total WTR Erasures by EOCT Subject

| Subject Area | N | Number of WTR Erasures | Mean | Standard Deviation | Correlation with Erasures | Number of Erasures by Percentiles | | | | | | Maximum Number of WTR Erasures |
|------------------|-------|------------------------|-------|--------------------|---------------------------|-----------------------------------|----|----|----|----|------|--------------------------------|
| | | | | | | 50 | 75 | 90 | 95 | 99 | 99.9 | |
| Mathematics I | 74846 | 15964 | 0.213 | 0.632 | 0.799 | 0 | 0 | 1 | 1 | 3 | 6 | 20 |
| Mathematics II | 70099 | 16426 | 0.234 | 0.671 | 0.805 | 0 | 0 | 1 | 1 | 3 | 6 | 15 |
| GPS Algebra | 18350 | 8061 | 0.439 | 0.914 | 0.817 | 0 | 1 | 1 | 2 | 4 | 7 | 11 |
| GPS Geometry | 7372 | 3176 | 0.431 | 0.963 | 0.844 | 0 | 1 | 1 | 2 | 5 | 8 | 11 |
| US History | 57705 | 18958 | 0.329 | 0.848 | 0.844 | 0 | 0 | 1 | 2 | 4 | 8 | 18 |
| Economics | 29370 | 12214 | 0.416 | 0.970 | 0.844 | 0 | 0 | 1 | 2 | 4 | 9 | 28 |
| Biology | 68259 | 23995 | 0.352 | 0.832 | 0.829 | 0 | 0 | 1 | 2 | 4 | 7 | 16 |
| Physical Science | 44909 | 19046 | 0.424 | 0.914 | 0.827 | 0 | 1 | 1 | 2 | 4 | 8 | 18 |
| 9th Grade Lit. | 76001 | 25899 | 0.341 | 0.836 | 0.848 | 0 | 0 | 1 | 2 | 4 | 7 | 26 |
| American Lit. | 59837 | 24770 | 0.414 | 0.929 | 0.852 | 0 | 1 | 1 | 2 | 4 | 8 | 24 |

Number of Schools Flagged for Erasure and WTR Analysis

Table 3 displays a summary of the number of schools flagged for **total erasures and wrong to right erasures based** on EOCT tests. The table includes subject area, total number of schools, number of schools flagged and % of schools flagged for erasures, and number of schools flagged and % of schools flagged for wrong to right.

Table 3. Number of Schools Flagged for Erasure and WTR Analysis

| Subject Area | Total Number of Schools | Erasures | | WTR | |
|------------------|-------------------------|---------------------------|----------------------|---------------------------|----------------------|
| | | Number of Schools Flagged | % of Schools Flagged | Number of Schools Flagged | % of Schools Flagged |
| Mathematics I | 490 | 68 | 13.88 | 60 | 12.24 |
| Mathematics II | 393 | 61 | 15.52 | 48 | 12.21 |
| GPS Algebra | 113 | 15 | 13.27 | 11 | 9.73 |
| GPS Geometry | 44 | 12 | 27.27 | 8 | 18.18 |
| US History | 378 | 73 | 19.31 | 57 | 15.08 |
| Economics | 308 | 44 | 14.29 | 35 | 11.36 |
| Biology | 397 | 86 | 21.66 | 64 | 16.12 |
| Physical Science | 385 | 55 | 14.29 | 40 | 10.39 |
| 9th Grade Lit. | 440 | 88 | 20.00 | 62 | 14.09 |
| American Lit. | 387 | 79 | 20.41 | 59 | 15.25 |

Number of Schools Flagged (WTR) in Any Subject Area for EOCT

Table 4 displays a summary of all schools with at least one class taking the EOCT for at least one subject. The table includes the following columns: total number of schools, number of schools flagged and % of schools flagged for wrong to right or changed to right responses, and number of schools not flagged and % of schools not flagged for wrong to right or changed to right responses, depending on the test delivery method. Table 4 shows that 66.27% of schools were not flagged in any subjects for EOCT tests.

Table 4. Number of Schools Flagged (WTR) in Any Subject Area for EOCT

| | Total Number of Schools | Number of Schools Flagged (WTR) | % of Schools Flagged (WTR) | Number of Schools Not Flagged (WTR) | % of Schools Not Flagged (WTR) |
|--------------------|-------------------------|---------------------------------|----------------------------|-------------------------------------|--------------------------------|
| Paper Tests | 593 | 200 | 33.73 | 393 | 66.27 |

Discussion

With the high-stakes nature of large-scale assessments such as the EOCT, there are times when examinee's scores may not be a true representation of his or her own abilities. This may occur due to an examinee copying from another examinee's paper, an examinee receiving inappropriate assistance before or during testing from a variety of sources, or an examinee's responses altered after testing. To maintain the validity of the EOCT results, it is important that occurrences, such as those previously mentioned, be discovered and identified.

It must be emphasized that the erasure analyses should only be considered as an initial step for checking a class with higher numbers of erasures than the state average. Flagging a class does not necessarily suggest improper activities. There are many potential sources of variances and alternative explanations are possible. Flagging should be an indicator to seek additional evidence to identify a possible problem within a class (and extended to a school or a district). Therefore, further investigation is imperative. The erasure analyses for Spring 2012 EOCT main administration is one component to uphold the integrity of the EOCT program and the assessment process.

These erasure analyses should only be used to identify *potential* problems within individual classrooms. These analyses must be confirmed by additional evidence before any conclusions regarding improprieties can be reached. In addition, when the class size is small, for example with 10 or fewer students, the erasure analysis results are only approximate and should be viewed with caution.