Spring 2016 Georgia Milestones Assessment Desktop Audit Results

February 2017

THE GOVERNOR'S OFFICE OF
STUDENT ACHIEVEMENT

Executive Summary

The Governor's Office of Student Achievement serves as the reporting and accountability agency for education in Georgia. As such, GOSA is charged by law with inspecting academic records of schools to ensure that education institutions are faithful to performance accountability requirements. Through an academic audit, GOSA reviews student assessment data and other school records reported to the State to confirm accuracy and explore the effectiveness of local school initiatives in improving achievement.

The Georgia Milestones Assessment Audit is divided into two separate analyses:

- **The Answer Change Analysis**, formerly known as the Erasure Analysis, identifies classrooms and schools where the number of wrong answers that have been changed to right answers on individual student answer sheets is well above the state average. For the first time, the 2015-2016 analysis includes assessments administered online. It is conducted on the following assessments:
    - Grades 3 to 8 End of Grade Assessments (EOG) in English-Language Arts, Mathematics, Science, and Social Studies
    - Grades 7-12 End of Course Assessments (EOC) in the following ten high school courses: Ninth Grade Literature and Composition, American Literature and Composition, Coordinate Algebra, Algebra I, Analytic Geometry, Geometry, Physical Science, Biology, U.S. History, and Economics.

- **The Unusual Response Pattern Analysis**, conducted for the first time this year, identifies schools that have unexpected test score gains across years using a cohort of students as well as unexpected patterns in student answers. It includes only EOG English and Mathematics in grades 4 to 8 and combines the following two indices:
    - Unexpected test score gains across years using a cohort of students (95^{th} percentile or higher), and
    - Unexpected patterns in student answers (95^{th} percentile or higher)
        - The patterns examined include (1) unlikely blocks of consecutive, identical answers, (2) highly correlated answers across tests, (3) correlation of responses across test items, and (4) cases where students miss easy items but answer difficult answers correctly.

Appendices C, D, and E provide more in-depth information on the calculation formulas and business rules. It is important to note that the results of both analyses are used as an initial flag to spur further investigation of many indicators to determine if any cheating occurred. The results do not indicate that cheating necessarily occurred.

Using the DRC Answer Change Analysis and Unusual Response Pattern Analysis, GOSA identifies schools for an internal desktop audit based on the following criteria:

Answer Change Analysis:

- EOG (Grades 3-8)
  - Five percent or more of classrooms in a school are flagged at four standard deviations or greater, OR
  - One classroom is flagged at seven standard deviations or greater, OR
- EOC (Grades 9-12)
  - Schools with multiple classrooms flagged at five standard deviations or greater, OR
  - One classroom is flagged at seven standard deviations or greater.

Unusual Response Pattern Analysis:

- EOG (Grades 4-8)[1]
  - Schools where two or more testing groups had test score gains and unusual response patterns that were in the 95th percentile, OR
  - Schools were one testing group had test score gains and unusual response patterns that were in the 99th percentile.

Results Summary

After the desktop audit, 108 schools in 45 LEAs have been identified for further inquiry, split out by analysis below:[2]

EOG Answer Change Analysis
222 classrooms in 75 schools in 40 LEAs were identified for an initial desktop audit. After the audit, 54 schools in 30 LEAs require further inquiry.

EOC Answer Change Analysis
76 classrooms in 30 schools in 12 LEAs were identified for an initial desktop audit. After the audit, 30 schools in 12 LEAs require further inquiry.

EOG Unusual Response Pattern Analysis
65 testing groups in 35 schools in 22 LEAs require further inquiry. A desktop audit was not conducted for the first year of this analysis.

Schools requiring further inquiry must conduct internal review and submit inquiry forms to GOSA by March 20, 2017. In addition, they must rotate teachers for the 2017 Georgia Milestones administration. State monitors will observe and inspect schools requiring further inquiry for the 2016 Georgia Milestones test administration (EOG and EOC tests). GOSA will conduct on-site audits as necessary.

The following report contains the results of GOSA's desktop audit and recommendations to the State Board of Education for actions to be taken in schools requiring further inquiry by the State.

[1] Each testing group is the total number of students by grade level and subject area (ELA or mathematics) who took a certain test form (A or B) regardless of classroom assignment. For example, all students in a school who took the 4th grade mathematics Georgia Milestones Form A assessment are a testing group.
[2] The number of schools flagged in each analysis does not add up to the total of 112 because seven schools were flagged in two of the analyses.

Table of Contents

Table of Abbreviations

GaDOE	Georgia Department of Education	LEA	Local Education Agency (District)
EOG	End of Grade Test	GOSA	Governor's Office of Student Achievement
EOC	End of Course Test	SBOE	State Board of Education
SD	Standard Deviation		

EOG Answer Change Desktop Audit Results

Analysis Overview

GOSA identified schools for a desktop audit when five percent or more of classrooms in a school were flagged at four standard deviations or greater, or one classroom was flagged at seven standard deviations or greater. In total, 222 classrooms in 75 schools in 40 LEAs were identified for an initial desktop audit to determine a possible explanation for the flag that would remove the need for further inquiry. In this audit analysis, many school-level factors, outlined in detail in Appendix A, were reviewed holistically and discussed as a team before any determinations were made. GOSA placed schools in one of two categories: "further inquiry needed," or "no further inquiry needed."

After the desktop audit, 54 schools in 30 LEAs, listed below, require further inquiry. Appendix B lists the number of classrooms at each school requiring further inquiry for both EOG and EOC.

Desktop Audit Schools

Schools Requiring Further Inquiry

- Atlanta Public Schools, Brandon Elementary School
- Atlanta Public Schools, Cleveland Elementary School
- Atlanta Public Schools, Hutchison Elementary School
- Atlanta Public Schools, Rivers Elementary School
- Bibb County, Alexander II Magnet School
- Bibb County, Brookdale Elementary School
- Bibb County, Porter Elementary School
- Bleckley County, Bleckley County Elementary School
- Camden County, Matilda Harris Elementary School
- Camden County, Saint Mary's Middle School
- Chatham County, Oglethorpe Charter School
- Chatham County, The Stem Academy at Bartlett
- Clarke County, Chase Street Elementary School
- Clinch County, Clinch County Elementary School
- Cobb County, Garrison Mill Elementary School
- Cobb County, Mount Bethel Elementary School
- Cobb County, Nickajack Elementary School
- Cobb County, Teasley Elementary School
- Cobb County, Tritt Elementary School
- Colquitt County, Norman Park Elementary School
- Commerce City, Commerce Middle School
- Coweta County, Brooks Elementary School
- Decatur City, Fifth Avenue Elementary School
- Decatur County, Jones-Wheat Elementary School
- DeKalb County, Cedar Grove Middle School
- DeKalb County, DeKalb Elementary School of the Arts
- DeKalb County, Indian Creek Elementary School
- DeKalb County, Peachtree Middle School

- DeKalb County, Vanderlyn Elementary School
- Dougherty County, Albany Middle School
- Dougherty County, Robert A Cross Middle Magnet School
- Fayette County, Crabapple Lane Elementary School
- Floyd County, Garden Lakes Elementary School
- Fulton County, Hembree Springs Elementary School
- Fulton County, Hopewell Middle School
- Gwinnett County, Gwinnett Online Campus
- Gwinnett County, Moore Middle School
- Gwinnett County, Mulberry Elementary School
- Gwinnett County, North Gwinnett Middle School
- Gwinnett County, Partee Elementary School
- Henry County, East Lake Elementary School
- Marion County, Marion County Middle-High School
- Murray County, Coker Elementary School
- Muscogee County, Britt David Elementary Computer Magnet Academy
- Muscogee County, Veterans Memorial Middle School
- Oconee County, High Shoals Elementary School
- Pierce County, Blackshear Elementary School
- Polk County, Eastside Elementary School
- Richmond County, Gracewood Elementary School
- Rockdale County, Shoal Creek Elementary School
- Tift County, G.O. Bailey Primary School
- Tift County, Len Lastinger Primary School
- Warren County, Freeman Elementary School
- Whitfield County, Cohutta Elementary School

EOC Answer Change Desktop Audit Results

Schools were identified for a desktop audit when multiple classrooms were flagged at five standard deviations or greater, or one classroom was flagged at seven standard deviations or greater. In total, 76 classrooms in 30 schools in 12 LEAs were identified for an initial desktop audit to determine a possible explanation for the flag that would remove the need for further inquiry. In this audit analysis, many school-level factors, outlined in detail in Appendix A, were reviewed holistically and discussed as a team before any determinations were made. GOSA placed schools in one of two categories: "further inquiry needed," or "no further inquiry needed."

After the desktop audit, 72 classrooms in 30 schools in 12 LEAs, listed below, require further inquiry. Appendix B lists the number of classrooms at each school requiring further inquiry for both EOG and EOC.

EOC Desktop Audit Schools

Schools Requiring Further Inquiry

- Cherokee County, River Ridge High School
- Columbia County, Lakeside High School
- DeKalb County, Arabia Mountain High School
- DeKalb County, Chamblee Charter High School
- DeKalb County, Chamblee Middle School
- DeKalb County, Clarkston High School
- DeKalb County, Cross Keys High School
- DeKalb County, Peachtree Middle School
- Effingham County, South Effingham High School
- Fayette County, McIntosh High School
- Fayette County, Whitewater High School
- Forsyth County, Lambert High School
- Forsyth County, North Forsyth High School
- Forsyth County, Riverwatch Middle School
- Forsyth County, South Forsyth High School
- Forsyth County, South Forsyth Middle School
- Fulton County, Alpharetta High School
- Fulton County, Autrey Mill Middle School
- Fulton County, Johns Creek High School
- Fulton County, Northview High School
- Fulton County, River Trail Middle School
- Fulton County, Webb Bridge Middle School
- Hall County, CW Davis Middle School
- Henry County, Luella High School
- Henry County, Union Grove High School
- Muscogee County, Columbus High School
- Muscogee County, Jordan Vocational High School
- Muscogee County, Northside High School
- Paulding County, Sammy McClure, Sr Middle School
- Peach County, Peach County High School

EOG Unusual Response Pattern Desktop Audit Results

Analysis Overview

GOSA identified schools for further inquiry that met one of the following criteria:

- Schools where two or more testing groups had test score gains and unusual response patterns that were in the 95th percentile, OR
- Schools were one testing group had test score gains and unusual response patterns that were in the 99th percentile.

Each testing group is the total number of students by grade level and subject area (ELA or mathematics) who took a certain test form (A or B) regardless of classroom assignment. For example, all students in a school who took the 4th grade mathematics Georgia Milestones Form A assessment are a testing group.

Since this is the first year of the analysis, all schools identified under the criteria are kept for further inquiry.

Using these criteria, 35 schools in 22 LEAs, listed below, require further inquiry. Appendix B lists the subject areas, test forms, and grade levels flagged for each school.

Desktop Audit Schools

Schools Requiring Further Inquiry

- Appling County, Appling County Middle School
- Bibb County, Skyview Elementary School
- Brooks County, Brooks County Middle School
- Buford City, Buford Middle School
- Carroll County, Bay Springs Middle School
- Catoosa County, Ringgold Middle School
- Cobb County, McClure Middle School
- Decatur City, Fifth Avenue Elementary School
- DeKalb County, Cary Reynolds Elementary School
- DeKalb County, Indian Creek Elementary School
- DeKalb County, Woodward Elementary School
- Early County, Early County Elementary School
- Fayette County, Rising Starr Middle School
- Fayette County, Whitewater Middle School
- Forsyth County, Brookwood Elementary School
- Forsyth County, George W. Whitlow Elementary School
- Forsyth County, Johns Creek Elementary School
- Forsyth County, North Forsyth Middle School
- Forsyth County, Silver City Elementary School
- Forsyth County, South Forsyth Middle School
- Gilmer County, Clear Creek Middle School
- Gwinnett County, Hull Middle School
- Gwinnett County, North Gwinnett Middle School

- Gwinnett County, Simpson Elementary School
- Gwinnett County, White Oak Elementary School
- Hall County, C.W. Davis Middle School
- Hall County, West Hall Middle School
- Heard County, Heard County Middle School
- Madison County, Colbert Elementary School
- Madison County, Madison County Middle School
- Murray County, Gladden Middle School
- Oconee County, Malcom Bridge Middle School
- Oglethorpe County, Oglethorpe County Elementary School
- Warren County, Freeman Elementary School
- Wayne County, Odum Elementary School

GOSA Recommendations in Schools Requiring Further Inquiry

Overall, 112 schools in 49 LEAs require further inquiry. GOSA recommends the following actions to the SBOE in these schools:

GOSA will:

1. Share EOG/EOC data files with superintendents of LEAs that have schools requiring further inquiry to facilitate:
    - LEA internal investigation of reason(s) for flags, and
    - Submission of online inquiry form to GOSA with results of investigation and an explanation of testing protocols in place.
2. Require identified schools to rotate teachers for the 2017 Georgia Milestones test administration (EOG).
3. Assign state monitors to observe and inspect identified schools requiring further inquiry for the 2017 Georgia Milestones test administration as necessary (EOG and EOG).
4. Conduct on-site audits as necessary.

No further action should be taken for flagged schools that require no further inquiry after the desktop audit.

Appendix A: Desktop Audit Indicators

Desktop Audit Indicators Reviewed
Number of classrooms flagged in each school and whether the flagged classrooms had different test administrators.
Total answer changes and number of wrong-to-right (w-t-r) at the classroom level, including student-level data to determine whether answer changes are concentrated in a small number of students. Classrooms where more than 50% of students in a classroom have zero answer changes and/or w-t-r answer changes reduce the likelihood of systematic or widespread changes in answers from wrong to right.
The severity of the individual flagged classroom (i.e. the standard deviation value or how far from what is considered normal behavior is the class positioned). EOC flags between 5.0 and 6.0 SDs are of less concern than those over 6.0 SDs.
Percentage of total classroom answer changes changed from w-t-r. Generally, classrooms with greater than 65% of answer changes being w-t-r are of concern, unless a classroom with multiple students had one student with many w-t-r answer changes, suggesting that systematic cheating was unlikely.
The number of students in each classroom. (Example: Extremes in classroom populations on both ends of the distribution can skew post-calculation metrics and in turn cause flagged classrooms.).
Classroom percentile ranks of wrong-to-right answer changes by student to observe the distribution of answer changes in a classroom and compare that distribution to the state distribution. For example, comparing a classroom's 50^{th} and 90^{th} percentile with the state 50^{th} and 90^{th} percentiles can identify whether abnormal distributions and/or outliers.
The type of school (i.e. high transient population, alternative education program, residential treatment facilities, etc.).
School demographics and groups (ELL population, gifted, magnet, students with disabilities, etc.).
Variance in performance level data from previous years (not applicable in 2014-2015 due to Georgia Milestones transition).
History as a school of concern.
Prior test monitoring and/or an on-site audit by state personnel.
District personnel and/or policies currently implemented to support test security.
Review of state monitor notes and/or forms.

Appendix B: EOG/EOC Schools Requiring Further Inquiry

The following list includes the number of classrooms flagged in the 112 schools requiring further inquiry after the desktop audit.

		2016 Spring EOG		2016 Spring EOC
System Name	School Name	Answer Change Classrooms Requiring Further Inquiry	Unusual Response Subject/Grade Form Requiring Further Inquiry	Answer Change Classrooms Requiring Further Inquiry
APPLING COUNTY	APPLING COUNTY MIDDLE SCHOOL		Math/6A, English/7B	
ATLANTA PUBLIC SCHOOLS	BRANDON ELEMENTARY SCHOOL	6		
ATLANTA PUBLIC SCHOOLS	CLEVELAND ELEMENTARY SCHOOL	8		
ATLANTA PUBLIC SCHOOLS	HUTCHINSON ELEMENTARY SCHOOL	4		
ATLANTA PUBLIC SCHOOLS	RIVERS ELEMENTARY SCHOOL	10		
BIBB COUNTY	ALEXANDER II MAGNET SCHOOL	2		
BIBB COUNTY	BROOKDALE ELEMENTARY SCHOOL	2		
BIBB COUNTY	PORTER ELEMENTARY SCHOOL	2		
BIBB COUNTY	SKYVIEW ELEMENTARY SCHOOL		English/4A, 4B	
BLECKLEY COUNTY	BLECKLEY COUNTY ELEMENTARY SCHOOL	1		
BROOKS COUNTY	BROOKS COUNTY MIDDLE SCHOOL		Math/8A, 8B	
BUFORD CITY	BUFORD MIDDLE SCHOOL		Math/6A, 6B	
CAMDEN COUNTY	MATILDA HARRIS ELEMENTARY SCHOOL	4		
CAMDEN COUNTY	SAINT MARYS MIDDLE SCHOOL	12		
CARROLL COUNTY	BAY SPRINGS MIDDLE SCHOOL		Math/8A, 8B	
CATOOSA COUNTY	RINGGOLD MIDDLE SCHOOL		Math/8A, 8B	
CHATHAM COUNTY	OGLETHORPE CHARTER SCHOOL	3		
CHATHAM COUNTY	THE STEM ACADEMY AT BARTLETT	1		
CHEROKEE COUNTY	RIVER RIDGE HIGH SCHOOL			2
CLARKE COUNTY	CHASE STREET ELEMENTARY SCHOOL	1		
CLINCH COUNTY	CLINCH COUNTY ELEMENTARY SCHOOL	3		
COBB COUNTY	GARRISON MILL ELEMENTARY SCHOOL	1		
COBB COUNTY	MCCLURE MIDDLE SCHOOL		Math/6A, 6B	
COBB COUNTY	MOUNT BETHEL ELEMENTARY SCHOOL	11		
COBB COUNTY	NICKAJACK ELEMENTARY SCHOOL	7		
COBB COUNTY	TEASLEY ELEMENTARY SCHOOL	7		
COBB COUNTY	TRITT ELEMENTARY SCHOOL	7		
COLQUITT COUNTY	NORMAN PARK ELEMENTARY SCHOOL	1		
COLUMBIA COUNTY	LAKESIDE HIGH SCHOOL			2

		2016 Spring EOG		2016 Spring EOC
System Name	**School Name**	**Answer Change Classrooms Requiring Further Inquiry**	**Unusual Response Subject/Grade Form Requiring Further Inquiry**	**Answer Change Classrooms Requiring Further Inquiry**
COMMERCE CITY	COMMERCE MIDDLE SCHOOL	3		
COWETA COUNTY	BROOKS ELEMENTARY SCHOOL	1		
DECATUR CITY	FIFTH AVENUE ELEMENTARY SCHOOL	12	English/5A	
DECATUR COUNTY	JONES-WHEAT ELEMENTARY SCHOOL	3		
DEKALB COUNTY	ARABIA MOUNTAIN HIGH SCHOOL			5
DEKALB COUNTY	CARY REYNOLDS ELEMENTARY SCHOOL		English/4A	
DEKALB COUNTY	CEDAR GROVE MIDDLE SCHOOL	1		
DEKALB COUNTY	CHAMBLEE CHARTER HIGH SCHOOL			2
DEKALB COUNTY	CHAMBLEE MIDDLE SCHOOL			1
DEKALB COUNTY	CLARKSTON HIGH SCHOOL			4
DEKALB COUNTY	CROSS KEYS HIGH SCHOOL			2
DEKALB COUNTY	DEKALB ELEMENTARY SCHOOL OF THE ARTS	6		
DEKALB COUNTY	INDIAN CREEK ELEMENTARY SCHOOL	1	Math/5A	
DEKALB COUNTY	PEACHTREE MIDDLE SCHOOL	17		2
DEKALB COUNTY	VANDERLYN ELEMENTARY SCHOOL	1		
DEKALB COUNTY	WOODWARD ELEMENTARY SCHOOL		Math/5A, 5B	
DOUGHERTY COUNTY	ALBANY MIDDLE SCHOOL	4		
DOUGHERTY COUNTY	ROBERT A. CROSS MIDDLE MAGNET	6		
EARLY COUNTY	EARLY COUNTY ELEMENTARY SCHOOL		Math/5A, 5B	
EFFINGHAM COUNTY	SOUTH EFFINGHAM HIGH SCHOOL			1
FAYETTE COUNTY	CRABAPPLE LANE ELEMENTARY SCHOOL	1		
FAYETTE COUNTY	MCINTOSH HIGH SCHOOL			2
FAYETTE COUNTY	RISING STARR MIDDLE SCHOOL		Math/7B, 8A, 8B	
FAYETTE COUNTY	WHITEWATER HIGH SCHOOL			1
FAYETTE COUNTY	WHITEWATER MIDDLE SCHOOL		English/7A, 7B, 8A, 8B	
FLOYD COUNTY	GARDEN LAKES ELEMENTARY SCHOOL	3		
FORSYTH COUNTY	BROOKWOOD ELEMENTARY SCHOOL		Math/4A, 4B	
FORSYTH COUNTY	GEORGE W. WHITLOW ELEMENTARY SCHOOL		English/4A, 4B	
FORSYTH COUNTY	JOHNS CREEK ELEMENTARY SCHOOL		English/5A	
FORSYTH COUNTY	LAMBERT HIGH SCHOOL			5
FORSYTH COUNTY	NORTH FORSYTH HIGH SCHOOL			2
FORSYTH COUNTY	NORTH FORSYTH MIDDLE SCHOOL		English/6A, 6B	
FORSYTH COUNTY	RIVERWATCH MIDDLE SCHOOL			4
FORSYTH COUNTY	SILVER CITY ELEMENTARY SCHOOL		Math/4A, 4B	
FORSYTH COUNTY	SOUTH FORSYTH HIGH SCHOOL			6

		2016 Spring EOG		2016 Spring EOC
System Name	**School Name**	**Answer Change Classrooms Requiring Further Inquiry**	**Unusual Response Subject/Grade Form Requiring Further Inquiry**	**Answer Change Classrooms Requiring Further Inquiry**
FORSYTH COUNTY	SOUTH FORSYTH MIDDLE SCHOOL		English/6A	3
FULTON COUNTY	ALPHARETTA HIGH SCHOOL			3
FULTON COUNTY	AUTREY MILL MIDDLE SCHOOL			1
FULTON COUNTY	HEMBREE SPRINGS ELEMENTARY SCHOOL	2		
FULTON COUNTY	HOPEWELL MIDDLE SCHOOL	3		
FULTON COUNTY	JOHNS CREEK HIGH SCHOOL			2
FULTON COUNTY	NORTHVIEW HIGH SCHOOL			3
FULTON COUNTY	RIVER TRAIL MIDDLE SCHOOL			3
FULTON COUNTY	WEBB BRIDGE MIDDLE SCHOOL			1
GILMER COUNTY	CLEAR CREEK MIDDLE SCHOOL		Math/8A, 8B	
GWINNETT COUNTY	GWINNETT ONLINE CAMPUS	3		
GWINNETT COUNTY	HULL MIDDLE SCHOOL		Math/7A, 7B	
GWINNETT COUNTY	MOORE MIDDLE SCHOOL	1		
GWINNETT COUNTY	MULBERRY ELEMENTARY SCHOOL	3		
GWINNETT COUNTY	NORTH GWINNETT MIDDLE SCHOOL	1	English/6A	
GWINNETT COUNTY	PARTEE ELEMENTARY SCHOOL	1		
GWINNETT COUNTY	SIMPSON ELEMENTARY SCHOOL		Math/4A, 4B	
GWINNETT COUNTY	WHITE OAK ELEMENTARY SCHOOL		English/4A, Math/4B	
HALL COUNTY	C. W. DAVIS MIDDLE SCHOOL		Math/6A, 6B	2
HALL COUNTY	WEST HALL MIDDLE SCHOOL		English/6B; Math/6A, 6B	
HEARD COUNTY	HEARD COUNTY MIDDLE SCHOOL		English/8A, Math/8A	
HENRY COUNTY	EAST LAKE ELEMENTARY SCHOOL	1		
HENRY COUNTY	LUELLA HIGH SCHOOL			2
HENRY COUNTY	UNION GROVE HIGH SCHOOL			1
MADISON COUNTY	COLBERT ELEMENTARY SCHOOL		Math/5A, 5B	
MADISON COUNTY	MADISON COUNTY MIDDLE SCHOOL		Math/7A, 7B	
MARION COUNTY	MARION COUNTY MIDDLE - HIGH SCHOOL	1		
MURRAY COUNTY	COKER ELEMENTARY SCHOOL	1		
MURRAY COUNTY	GLADDEN MIDDLE SCHOOL		Math/8A, 8B	
MUSCOGEE COUNTY	BRITT DAVID ELEMENTARY COMPUTER MAGNET	4		
MUSCOGEE COUNTY	COLUMBUS HIGH SCHOOL			4
MUSCOGEE COUNTY	JORDAN VOCATIONAL HIGH SCHOOL			1
MUSCOGEE COUNTY	NORTHSIDE HIGH SCHOOL			2
MUSCOGEE COUNTY	VETERANS MEMORIAL MIDDLE SCHOOL	3		

		2016 Spring EOG		2016 Spring EOC
System Name	**School Name**	**Answer Change Classrooms Requiring Further Inquiry**	**Unusual Response Subject/Grade Form Requiring Further Inquiry**	**Answer Change Classrooms Requiring Further Inquiry**
OCONEE COUNTY	HIGH SHOALS ELEMENTARY SCHOOL	2		
OCONEE COUNTY	MALCOM BRIDGE MIDDLE SCHOOL		Math/7A, 7B	
OGLETHORPE COUNTY	OGLETHORPE COUNTY ELEMENTARY SCHOOL		Math/4B	
PAULDING COUNTY	SAMMY MCCLURE SR MIDDLE SCHOOL			2
PEACH COUNTY	PEACH COUNTY HIGH SCHOOL			1
PIERCE COUNTY	BLACKSHEAR ELEMENTARY SCHOOL	1		
POLK COUNTY	EASTSIDE ELEMENTARY SCHOOL	4		
RICHMOND COUNTY	GRACEWOOD ELEMENTARY SCHOOL	1		
ROCKDALE COUNTY	SHOAL CREEK ELEMENTARY SCHOOL	5		
TIFT COUNTY	G. O. BAILEY PRIMARY SCHOOL	2		
TIFT COUNTY	LEN LASTINGER PRIMARY SCHOOL	3		
WARREN COUNTY	FREEMAN ELEMENTARY SCHOOL	3	English/4A	
WAYNE COUNTY	ODUM ELEMENTARY SCHOOL		Math/4A	
WHITFIELD COUNTY	COHUTTA ELEMENTARY SCHOOL	2		
Totals:	**112 schools**	**199 classrooms**	**65 Form/Grade/ Subject Area Combinations**	**72 classrooms**

Appendix C: 2016 DRC EOG Answer Change Executive Report

Analysis of Answer Changes

Submitted by DRC

January 2017

With the high-stakes nature of large-scale assessments such as the Milestones End of Grade (EOG), there are times when students' responses, and hence their scores, may not be a true representation of their own abilities. Various activities may take place, such as a student copying from another student's paper, students receiving inappropriate assistance before or during testing, or students' responses altered after testing. To maintain the integrity of the Milestones EOG and the validity of the results, it is important that any such instances be discovered.

The present study investigated student responses on the English Language/Arts, Mathematics, Science and Social Studies tests of the 2016 Spring Milestones EOG that a) an answer choice was replaced by a different answer choice and b) changed from a wrong answer to a right answer (wrong-to-right).

It should be emphasized that results from the erasure analyses performed in 2016 should only be used to identify potential problems within individual classrooms. That is, these types of analyses must be supported by additional, collateral information before conclusions regarding any improprieties are reached.

Answer Changes for Paper Administrations

The GA Milestones EOG paper-pencil answer documents were processed using high speed 5000i optical scanners which reliably captured document images and optical mark read data. The sophisticated proprietary scoring software system, specifically Optical Mark Recognition (OMR) software, reviews the integrity of each batch of documents scanned according to pre-defined guidelines and services.

The OMR software provides a mechanism for identifying multiple-marks and identification of erasures for scanned data to support answer change analysis. The basis of the answer change analysis is to count erasures for multiple-choice items where two or more responses have been made with a specified intensity. Erasure analyses provide a mechanism to differentiate between three kinds of answer changes: a) wrong-to-wrong, b) right-to-wrong and c) wrong-to-right. Capturing the frequency of answer changes from wrong-to-right can be useful for identifying potential instances of cheating at the student level. Erasure analyses results can be grouped to tentatively identify problems at the classroom and school levels.

Answer Changes for Online Administrations

The test administration software that delivers the Georgia Milestones assessment system, INSIGHT, captures answer changes during online testing sessions. Similar to paper based administrations where answer changes are determined by examining erasure marks, the INSIGHT system records changes to answers within an online test administration that are made either before leaving an item or upon returning to the item and making a change. Answer change analyses for

students testing online also focuses on the three kinds of changes: a) wrong-to-wrong, b) right-to-wrong and c) wrong-to-right. As with paper based erasure analyses, capturing the frequency of answer changes from wrong-to-right can be useful for identifying potential instances of cheating at the student level in online testing. Analyses results can be grouped to tentatively identify problems at the classroom and school levels.

Method

The basis for the answer change analysis is to count erasures in items where an answer choice was erased and replaced with another answer choice; online, an item was selected and then later changed to a different answer choice. Herein, both actions are referred to as an erasure. Often the data captured is useful for identifying cases of cheating. During erasure analysis, two sets of erasures were analyzed: all erasures and wrong-to-right erasures where an incorrect answer choice was erased and replaced with the correct answer choice. Only operational items were used for the answer change analyses implemented for the 2016 Georgia Milestones.

The basic idea underlying the procedure is a statistical test of the null hypothesis (H_0) that the mean number of erasures for a class constitutes a random sample from the state distribution of erasures. The hypothesis is tested against the (right-sided) alternative (H_1) that the mean number is too high to be explained by random sampling. Classes for which H_0 has to be rejected are flagged for further scrutiny. A well-known central limit theorem in statistics tells us that the sampling distribution of the mean number of erasures for class i (m_i) is asymptotically normal with mean and standard deviation (SD)

$$mean(m_i) = \mu \qquad (1)$$

$$SD(m_i) = \frac{\sigma}{\sqrt{n_i}} \qquad (2)$$

where n_i and m_i denote the size and mean number of erasures for class i, respectively. In addition, μ and σ denote the mean and the SD of the distribution of the number of erasures of the population of individual students in the state of Georgia.

The classes were flagged if their m_i was larger than $\mu + 4\frac{\sigma}{\sqrt{n_i}}$. Statistically, the flagging criterion set at or above 4σ is conservative. The standard normal table shows that under random sampling the (asymptotic) probability of a sample mean being more than four SDs above the population mean is around 0.00003. However, rejection of H_0 only tells us that the observed mean number of erasures is unlikely to be the result of random sampling.

It is evident in the formula that the class flagging criterion for each class is adjusted for the number of test takers in a classroom. For example, if the state mean and SD of erasure count are 1.73 and 2.11, respectively, the flagging criterion for a class size of 20 is adjusted to 3.62 ($1.73 + 4\frac{2.11}{\sqrt{20}} = 3.62$).

This adjustment ensures that the flagging criterion is equally stringent for classes with considerably different numbers of test takers. In addition, minimizing the probability of false positive (Type I) errors in this statistical test is crucial in this analysis.

Results

Tables 1 and 2 reports the state summary of erasure counts for paper-pencil and online respectively. The tables include the number of students, the total number of all types of erasures, the mean and the SD of all types of erasures, the correlation between all erasures and wrong-to-right erasures, the number of erasures at the 50th, 75th, 90th, 95th, 99th, and 99.9th percentiles, and the maximum number of all types of erasures. The mean number of paper-pencil erasures across all courses ranged from 0.69 to 1.47, and mean number of online answer changes ranged from 5.97 to 9.72 for the 2016 Spring Milestones EOG. In other words, approximately 1 to 2 answer changes were made per student paper-pencil answer sheet on average, and 6 to 10 answer changes were made per online student assessment. The erasure count at specific percentile points (50th, 75th, 90th, 95th, 99th, and 99.9th) is also reported. The erasure count at the 95th percentile point was between 3 and 5 on paper-pencil answer sheets, and between 13 and 21 online.

Tables 3 and 4 report the state summary of wrong-to-right erasure counts for paper-pencil and online respectively. The tables include the number of students, the number of wrong-to-right erasures, the mean and the SD of wrong-to-right erasures, the correlation between all erasures and wrong-to-right erasures, the number of wrong-to-right erasure at the 50th, 75th, 90th, 95th, 99th, and 99.9th percentiles, and the maximum number of wrong-to-right erasures. As can be expected, the mean wrong-to-right erasure count and the count at the specific percentile points were lower than those obtained from all erasure counts. The mean number of paper-pencil wrong-to-right erasures ranged from 0.36 to 0.84, and mean number of online answer changes ranged from 2.55 to 4.55 for the 2016 Spring Milestones EOG. In other words, approximately 0 to 1 wrong-to-right answer changes were made per student paper-pencil answer sheet on average, and approximately 3 to 5 wrong-to-right answer changes were made per online student. The wrong-to-right erasure count at specific percentile points (50th, 75th, 90th, 95th, 99th, and 99.9th) is also reported. The wrong-to-right erasure count at the 95th percentile point was between 2 and 3 on paper-pencil answer sheets, and between 6 and 11 online.

Tables 5 and 6 present a summary of the number of schools flagged across four content areas - English Language/Arts, Mathematics, Science and Social Studies – within each analysis of the Milestones EOG for paper-pencil and online respectively. For each analysis, the number of schools was computed in two ways: flagged for at least one content area or flagged for all four content areas. The number/percentage of schools that had zero flags for all erasures and wrong-to-right erasures in English Language/Arts, Mathematics, Science and Social Studies is provided in Tables 7 and 8. The number/percentage of schools that had less than 1% of the classes flagged for all erasures and wrong-to-right erasures in English Language/Arts, Mathematics, Science and Social Studies and across grades is provided in Tables 9 and 10.

Discussion

With respect to the erasure analyses, the following caveats are always applicable:

1. The normal distribution holds only for large classes; for smaller classes the result is approximate.

2. Rejection of H_0 does not necessarily imply cheating. Alternative explanations are

possible.

3. The flagging criterion should thus be taken as a stimulus to look for additional evidence and find out what happened in the school.

4. The groups of students taking the tests online and paper are not equivalent. Comparing the magnitude of answer changes between testing modes cannot be supported given that the groups likely differ in ability and other key background characteristic.

This answer change analysis is considered a check for unusual numbers of answer changes to student responses. Without additional layers added to the analysis, this kind of check only addresses the possibility, not the certainty, of teachers or administrators altering the responses of students. The 2016 erasure analyses represent an important step in helping to maintain the integrity of future administrations of the Milestones EOG.

Table 1. State Summary Statistics for All Types of Erasure (ERA) Counts by Content/Grade Paper-Pencil

Content	Grade	N	No. of Erasures	Mean	SD	Correlation between ERA and WTR	Number of Erasure by Percentiles						Max
							50	75	90	95	99	99.9	
ELA	3	73900	83579	1.13	1.61	0.78	1	2	3	4	7	12	46
	4	62936	43137	0.69	1.23	0.80	0	1	2	3	5	9	46
	5	37017	30643	0.83	1.35	0.80	0	1	2	3	6	10	45
	6	73585	60045	0.82	1.31	0.81	0	1	2	3	6	10	47
	7	64339	50099	0.78	1.26	0.80	0	1	2	3	6	10	34
	8	39272	36232	0.92	1.36	0.83	0	1	3	4	6	10	17
MA	3	73900	82694	1.12	1.68	0.82	1	2	3	4	7	13	35
	4	62936	51582	0.82	1.42	0.82	0	1	2	4	6	11	57
	5	37017	32641	0.88	1.42	0.82	0	1	3	4	6	10	21
	6	73585	63380	0.86	1.39	0.82	0	1	3	4	6	10	47
	7	64339	56446	0.88	1.41	0.80	0	1	3	4	6	11	29
	8	39272	33687	0.86	1.46	0.81	0	1	3	4	6	10	21
SC	3	73900	100820	1.36	2.04	0.85	1	2	4	5	9	17	37
	4	62936	54124	0.86	1.60	0.84	0	1	3	4	7	13	64
	5	37017	36051	0.97	1.66	0.85	0	1	3	4	7	12	30
	6	73585	73926	1.00	1.66	0.84	0	1	3	4	7	13	31
	7	64339	71829	1.12	1.76	0.84	0	2	3	4	8	15	40
	8	39272	32573	0.83	1.63	0.85	0	1	3	4	7	14	31
SS	3	73900	108695	1.47	2.30	0.87	1	2	4	5	10	22	52
	4	62936	49969	0.79	1.58	0.85	0	1	2	4	7	14	57
	5	37017	38903	1.05	1.82	0.86	0	1	3	4	8	17	31
	6	73585	67293	0.91	1.68	0.86	0	1	3	4	7	16	42
	7	64339	70839	1.10	1.87	0.85	0	2	3	4	8	18	52
	8	39272	50585	1.29	2.00	0.86	1	2	3	5	8	18	59

5

Table 2. State Summary Statistics for All Types of Erasure (ERA) Counts by Content/Grade Online

Content	Grade	N	No. of Erasures	Mean	SD	Correlation between ERA and WTR	Number of Erasure by Percentiles						Max
							50	75	90	95	99	99.9	
ELA	3	62267	454164	7.29	4.57	0.68	7	10	13	16	22	31	48
	4	69702	448057	6.43	4.25	0.69	6	9	12	14	20	28	48
	5	92946	628485	6.76	4.26	0.70	6	9	12	15	20	28	48
	6	55132	355244	6.44	4.15	0.72	6	9	12	14	19	28	44
	7	63697	416339	6.54	4.20	0.75	6	9	12	14	20	27	39
	8	89853	574549	6.39	4.16	0.77	6	9	12	14	20	27	50
MA	3	62489	404775	6.48	4.29	0.67	6	9	12	14	20	32	62
	4	69944	433535	6.20	4.07	0.64	5	8	11	14	19	29	61
	5	93278	577583	6.19	4.11	0.65	5	8	11	14	19	29	53
	6	55308	330137	5.97	3.91	0.67	5	8	11	13	18	27	49
	7	63776	418082	6.56	4.24	0.67	6	9	12	14	20	29	54
	8	86679	584842	6.75	4.76	0.73	6	9	13	15	21	29	51
SC	3	62411	489069	7.84	5.08	0.74	7	10	14	17	24	38	65
	4	69865	593469	8.49	5.14	0.72	8	11	15	18	25	36	63
	5	93170	790331	8.48	5.11	0.75	8	11	15	18	25	35	64
	6	55192	477057	8.64	5.22	0.73	8	11	15	18	25	36	61
	7	63756	576302	9.04	5.59	0.76	8	12	16	19	27	39	63
	8	85808	717552	8.36	6.02	0.83	8	12	16	19	26	38	57
SS	3	62012	485812	7.83	4.98	0.74	7	10	14	17	24	37	63
	4	69468	580451	8.36	5.29	0.75	7	11	15	18	25	37	65
	5	92649	812471	8.77	5.38	0.75	8	12	16	19	26	39	66
	6	54840	482583	8.80	5.51	0.75	8	12	16	19	26	39	62
	7	63353	556239	8.78	5.46	0.76	8	12	16	19	26	39	65
	8	89373	868602	9.72	6.03	0.81	9	13	18	21	29	42	65

Table 3. State Summary Statistics for Wrong-to-Right (WTR) Erasure Counts by Content/Grade Paper-Pencil

Content	Grade	N	No. of Erasures	Mean	SD	Correlation between ERA and WTR	Number of Erasure by Percentiles						Max
							50	75	90	95	99	99.9	
ELA	3	73900	40798	0.55	0.94	0.78	0	1	2	2	4	7	24
	4	62936	22865	0.36	0.76	0.80	0	1	1	2	3	5	35
	5	37017	15813	0.43	0.82	0.80	0	1	1	2	4	6	18
	6	73585	33397	0.45	0.84	0.81	0	1	1	2	4	6	27
	7	64339	25800	0.40	0.77	0.80	0	1	1	2	3	5	12
	8	39272	20872	0.53	0.91	0.83	0	1	2	2	4	6	9
MA	3	73900	45775	0.62	1.08	0.82	0	1	2	3	5	8	25
	4	62936	26874	0.43	0.88	0.82	0	1	1	2	4	7	32
	5	37017	17862	0.48	0.91	0.82	0	1	2	2	4	7	10
	6	73585	35107	0.48	0.89	0.82	0	1	2	2	4	7	13
	7	64339	29209	0.45	0.87	0.80	0	1	1	2	4	7	15
	8	39272	16302	0.42	0.85	0.81	0	1	1	2	4	6	12
SC	3	73900	57869	0.78	1.35	0.85	0	1	2	3	6	11	24
	4	62936	28621	0.45	0.99	0.84	0	1	2	2	4	8	26
	5	37017	19896	0.54	1.07	0.85	0	1	2	3	5	9	14
	6	73585	39143	0.53	1.03	0.84	0	1	2	2	5	8	17
	7	64339	38382	0.60	1.12	0.84	0	1	2	3	5	10	22
	8	39272	15869	0.40	0.93	0.85	0	0	1	2	4	8	14
SS	3	73900	61909	0.84	1.50	0.87	0	1	2	3	7	14	41
	4	62936	28191	0.45	1.01	0.85	0	1	1	2	5	9	26
	5	37017	21727	0.59	1.19	0.86	0	1	2	3	5	12	25
	6	73585	37314	0.51	1.08	0.86	0	1	2	2	5	10	26
	7	64339	38109	0.59	1.18	0.85	0	1	2	3	5	11	41
	8	39272	27076	0.69	1.26	0.86	0	1	2	3	5	12	46

Table 4. State Summary Statistics for Wrong-to-Right (WTR) Erasure Counts by Content/Grade Online

Content	Grade	N	No. of Erasures	Mean	SD	Correlation between ERA and WTR	Number of Erasure by Percentiles						Max
							50	75	90	95	99	99.9	
ELA	3	62267	186700	3.00	2.30	0.68	3	4	6	7	10	15	24
	4	69702	190379	2.73	2.18	0.69	2	4	6	7	10	15	31
	5	92946	268762	2.89	2.28	0.70	2	4	6	7	10	15	31
	6	55132	161640	2.93	2.37	0.72	2	4	6	7	10	16	38
	7	63697	187862	2.95	2.41	0.75	2	4	6	7	11	16	31
	8	89853	287057	3.19	2.48	0.77	3	4	6	8	11	16	37
MA	3	62489	178175	2.85	2.17	0.67	2	4	6	7	10	14	40
	4	69944	182479	2.61	2.04	0.64	2	4	5	6	9	12	36
	5	93278	238315	2.55	2.00	0.65	2	4	5	6	9	13	31
	6	55308	141626	2.56	1.99	0.67	2	4	5	6	9	12	22
	7	63776	177936	2.79	2.15	0.67	2	4	6	7	9	14	46
	8	86679	235218	2.71	2.32	0.73	2	4	6	7	10	14	24
SC	3	62411	226287	3.63	2.76	0.74	3	5	7	9	12	20	41
	4	69865	261499	3.74	2.79	0.72	3	5	7	9	13	19	43
	5	93170	358007	3.84	2.81	0.75	3	5	7	9	13	19	49
	6	55192	201867	3.66	2.72	0.73	3	5	7	9	12	19	39
	7	63756	266713	4.18	3.17	0.76	4	6	8	10	14	22	53
	8	85808	323648	3.77	3.21	0.83	3	5	8	10	14	21	37
SS	3	62012	224190	3.62	2.69	0.74	3	5	7	8	12	19	44
	4	69468	263464	3.79	2.92	0.75	3	5	7	9	13	22	53
	5	92649	367145	3.96	2.98	0.75	3	5	8	9	14	22	42
	6	54840	220395	4.02	3.11	0.75	3	6	8	10	14	23	47
	7	63353	258186	4.08	3.11	0.76	4	6	8	10	14	22	50
	8	89373	406579	4.55	3.45	0.81	4	6	9	11	16	26	48

Table 5. Number of Schools Flagged for Erasure and WTR Erasure and WTR Analysis Paper-Pencil

Grade	Total Number of Schools	All Erasure Analyses		Wrong-to-Right Erasure Analyses	
		Number of Schools Flagged for at Least One Content Area	Number of Schools Flagged for All Content Areas	Number of Schools Flagged for at Least One Content Area	Number of Schools Flagged for All Content Areas
3	787	117	2	88	2
4	708	98	4	88	2
5	527	70	2	49	2
6	399	69	3	56	1
7	364	71	2	51	0
8	306	62	0	43	0

Table 6. Number of Schools Flagged for Erasure and WTR Erasure and WTR Analysis Online

Grade	Total Number of Schools	All Erasure Analyses		Wrong-to-Right Erasure Analyses	
		Number of Schools Flagged for at Least One Content Area	Number of Schools Flagged for All Content Areas	Number of Schools Flagged for at Least One Content Area	Number of Schools Flagged for All Content Areas
3	1051	139	4	62	0
4	1082	130	4	60	1
5	1184	123	3	64	0
6	542	68	6	37	0
7	529	48	3	50	0
8	553	88	7	63	2

Table 7. The number/percentage of schools that had zero flags for all erasures and wrong-to-right erasures Paper-Pencil

Grade	English/Language Arts			Mathematics			Science			Social Studies		
	No. of Schools	No. Of Schools with Zero Flags	% of Schools with Zero Flags	No. of Schools	No. Of Schools with Zero Flags	% of Schools with Zero Flags	No. of Schools	No. Of Schools with Zero Flags	% of Schools with Zero Flags	No. of Schools	No. Of Schools with Zero Flags	% of Schools with Zero Flags
3	787	728	93%	787	720	91%	787	730	93%	787	743	94%
4	708	662	94%	708	660	93%	708	663	94%	708	657	93%
5	527	495	94%	527	493	94%	527	493	94%	527	497	94%
6	399	357	89%	399	359	90%	399	364	91%	399	370	93%
7	364	329	90%	364	336	92%	364	334	92%	364	325	89%
8	306	280	92%	306	277	91%	306	268	88%	306	292	95%

Table 8. The number/percentage of schools that had zero flags for all erasures and wrong-to-right erasures Online

Grade	English/Language Arts			Mathematics			Science			Social Studies		
	No. of Schools	No. Of Schools with Zero Flags	% of Schools with Zero Flags	No. of Schools	No. Of Schools with Zero Flags	% of Schools with Zero Flags	No. of Schools	No. Of Schools with Zero Flags	% of Schools with Zero Flags	No. of Schools	No. Of Schools with Zero Flags	% of Schools with Zero Flags
3	1047	985	94%	1049	975	93%	1048	988	94%	1043	980	94%
4	1081	1004	93%	1080	1015	94%	1080	1017	94%	1079	1030	95%
5	1183	1121	95%	1183	1104	93%	1182	1126	95%	1181	1131	96%
6	542	504	93%	541	510	94%	541	508	94%	541	509	94%
7	528	493	93%	527	490	93%	528	498	94%	527	501	95%
8	553	503	91%	552	488	88%	542	493	91%	552	518	94%

10

Table 9. The number/percentage of schools that had less than 1% of the classes flagged for all erasures and wrong-to-right erasures Paper-Pencil

English/Language Arts			Mathematics			Science			Social Studies		
No. of Schools	No. of Schools with <1% Flag Across Grades	% of Schools with <1% Flag Across Grades	No. of Schools	No. of Schools with <1% Flag Across Grades	% of Schools with <1% Flag Across Grades	No. of Schools	No. of Schools with <1% Flag Across Grades	% of Schools with <1% Flag Across Grades	No. of Schools	No. of Schools with <1% Flag Across Grades	% of Schools with <1% Flag Across Grades
1395	1389	99.6%	1395	1392	99.8%	1395	1390	99.6%	1395	1392	99.8%

Table 10. The number/percentage of schools that had less than 1% of the classes flagged for all erasures and wrong-to-right erasures Online

English/Language Arts			Mathematics			Science			Social Studies		
No. of Schools	No. of Schools with <1% Flag Across Grades	% of Schools with <1% Flag Across Grades	No. of Schools	No. of Schools with <1% Flag Across Grades	% of Schools with <1% Flag Across Grades	No. of Schools	No. of Schools with <1% Flag Across Grades	% of Schools with <1% Flag Across Grades	No. of Schools	No. of Schools with <1% Flag Across Grades	% of Schools with <1% Flag Across Grades
1749	1748	99.9%	1749	1749	100.0%	1748	1748	100.0%	1748	1748	100.0%

Appendix D: 2016 DRC EOC Answer Change Executive Report

Analysis of Answer Changes

Submitted by DRC

January 2017

With the high-stakes nature of large-scale assessments such as the Milestones End of Course (EOC), there are times when students' responses, and hence their scores, may not be a true representation of their own abilities. Various activities may take place, such as a student copying from another student's paper, students receiving inappropriate assistance before or during testing, or students' responses altered after testing. To maintain the integrity of the Milestones EOC and the validity of the results, it is important that any such instances be discovered.

The present study investigated student responses on the Ninth Grade Literature & Composition, American Literature & Composition, Coordinate Algebra, Analytic Geometry, Algebra I, Geometry, Biology, Physical Science, United States History, and Economics tests of the 2016 Spring Milestones EOC that a) an answer choice was replaced by a different answer choice and b) changed from a wrong answer to a right answer (wrong-to-right).

It should be emphasized that results from the erasure analyses performed in 2016 should only be used to identify potential problems within individual classrooms. That is, these types of analyses must be supported by additional, collateral information before conclusions regarding any improprieties are reached.

Answer Changes for Paper Administrations

The GA Milestones EOC paper-pencil answer documents were processed using high speed 5000i optical scanners which reliably captured document images and optical mark read data. The sophisticated proprietary scoring software system, specifically Optical Mark Recognition (OMR) software, reviews the integrity of each batch of documents scanned according to pre-defined guidelines and services.

The OMR software provides a mechanism for identifying multiple-marks and identification of erasures for scanned data to support answer change analysis. The basis of the answer change analysis is to count erasures for multiple-choice items where two or more responses have been made with a specified intensity. Erasure analyses provide a mechanism to differentiate between three kinds of answer changes: a) wrong-to-wrong, b) right-to-wrong and c) wrong-to-right. Capturing the frequency of answer changes from wrong-to-right can be useful for identifying potential instances of cheating at the student level. Erasure analyses results can be grouped to tentatively identify problems at the classroom and school levels.

Answer Changes for Online Administrations

The test administration software that delivers the Georgia Milestones assessment system, INSIGHT, captures answer changes during online testing sessions. Similar to paper based administrations where answer changes are determined by examining erasure marks, the INSIGHT system records changes to answers within an online test administration that are made either before

leaving an item or upon returning to the item and making a change. Answer change analyses for students testing online also focuses on the three kinds of changes: a) wrong-to-wrong, b) right-to-wrong, and c) wrong-to-right. As with paper based erasure analyses, capturing the frequency of answer changes from wrong-to-right can be useful for identifying potential instances of cheating at the student level in online testing. Analyses results can be grouped to tentatively identify problems at the classroom and school levels.

Method

The basis for the answer change analysis is to count erasures in items where an answer choice was erased and replaced with another answer choice; online, an item was selected and then later changed to a different answer choice. Herein, both actions are referred to as an erasure. Often the data captured is useful for identifying cases of cheating. During erasure analysis, two sets of erasures were analyzed: all erasures and wrong-to-right erasures where an incorrect answer choice was erased and replaced with the correct answer choice. Only operational items were used for the answer change analyses implemented for the 2016 Georgia Milestones.

The basic idea underlying the procedure is a statistical test of the null hypothesis (H_0) that the mean number of erasures for a class constitutes a random sample from the state distribution of erasures. The hypothesis is tested against the (right-sided) alternative (H_1) that the mean number is too high to be explained by random sampling. Classes for which H_0 has to be rejected are flagged for further scrutiny. A well-known central limit theorem in statistics tells us that the sampling distribution of the mean number of erasures for class i (m_i) is asymptotically normal with mean and standard deviation (SD)

$$mean(m_i) = \mu \qquad (1)$$

$$SD(m_i) = \frac{\sigma}{\sqrt{n_i}} \qquad (2)$$

where n_i and m_i denote the size and mean number of erasures for class i, respectively. In addition, μ and σ denote the mean and the SD of the distribution of the number of erasures of the population of individual students in the state of Georgia.

The classes were flagged if their m_i was larger than $\mu + 5\frac{\sigma}{\sqrt{n_i}}$. Statistically, the flagging criterion set at or above 5σ is conservative. The standard normal table shows that under random sampling the (asymptotic) probability of a sample mean being more than five SDs above the population mean is around 0.00003. However, rejection of H_0 only tells us that the observed mean number of erasures is unlikely to be the result of random sampling.

It is evident in the formula that the class flagging criterion for each class is adjusted for the number of test takers in a classroom. For example, if the state mean and SD of erasure count are 1.73 and 2.11, respectively, the flagging criterion for a class size of 20 is adjusted to 4.11 ($1.73 + 5\frac{2.11}{\sqrt{20}} = 4.11$).

This adjustment ensures that the flagging criterion is equally stringent for classes with considerably different numbers of test takers. In addition, minimizing the probability of false

positive (Type I) errors in this statistical test is crucial in this analysis.

Results

Tables 1 and 2 reports the state summary of erasure counts for paper-pencil and online respectively. The tables include the number of students, the total number of all types of erasures, the mean and the SD of all types of erasures, the correlation between all erasures and wrong-to-right erasures, the number of erasures at the 50^{th}, 75^{th}, 90^{th}, 95^{th}, 99^{th}, and 99.9^{th} percentiles, and the maximum number of all types of erasures. The mean number of paper-pencil erasures across all courses ranged from 0.63 to 1.31, and mean number of online answer changes ranged from 4.47 to 9.51 for the 2016 Spring Milestones EOC. In other words, approximately 1 answer change was made per student paper-pencil answer sheet on average, and 4 to 10 answer changes were made per online student assessment. The erasure count at specific percentile points (50^{th}, 75^{th}, 90^{th}, 95^{th}, 99^{th}, and 99.9^{th}) is also reported. The erasure count at the 95^{th} percentile point was between 3 and 5 on paper-pencil answer sheets, and between 11 and 21 online.

Tables 3 and 4 report the state summary of wrong-to-right erasure counts for paper-pencil and online respectively. The tables include the number of students, the number of wrong-to-right erasures, the mean and the SD of wrong-to-right erasures, the correlation between all erasures and wrong-to-right erasures, the number of wrong-to-right erasure at the 50^{th}, 75^{th}, 90^{th}, 95^{th}, 99^{th}, and 99.9^{th} percentiles, and the maximum number of wrong-to-right erasures. As can be expected, the mean wrong-to-right erasure count and the count at the specific percentile points were lower than those obtained from all erasure counts. The mean number of paper-pencil wrong-to-right erasures ranged from 0.35 to 0.59, and mean number of online answer changes ranged from 2.49 to 4.55 for the 2016 Spring Milestones EOC. In other words, approximately 0 to 1 wrong-to-right answer changes were made per student paper-pencil answer sheet on average, and approximately 2 to 5 wrong-to-right answer changes were made per online student. The wrong-to-right erasure count at specific percentile points (50^{th}, 75^{th}, 90^{th}, 95^{th}, 99^{th}, and 99.9^{th}) is also reported. The wrong-to-right erasure count at the 95^{th} percentile point was between 2 and 3 on paper-pencil answer sheets, and between 7 and 11 online.

Tables 5 and 6 present a summary of the number of schools flagged for total erasures and wrong-to-right erasures based on Milestones EOC paper-pencil and online respectively. Table 7 presents a summary of all schools with at least one class taking the Milestones EOC for at least one subject.

Discussion

With respect to the erasure analyses, the following caveats are always applicable:

1. The normal distribution holds only for large classes; for smaller classes the result is approximate.

2. Rejection of H_0 does not necessarily imply cheating. Alternative explanations are possible.

3. The flagging criterion should thus be taken as a stimulus to look for additional evidence and find out what happened in the school.

4. The groups of students taking the tests online and paper are not equivalent. Comparing the magnitude of answer changes between testing modes cannot be supported given that the groups likely differ in ability and other key background characteristic.

This answer change analysis is considered a check for unusual numbers of answer changes to student responses. Without additional layers added to the analysis, this kind of check only addresses the possibility, not the certainty, of teachers or administrators altering the responses of students. The 2016 erasure analyses represent an important step in helping to maintain the integrity of future administrations of the Milestones EOC.

Table 1. State Summary Statistics for All Types of Erasure (ERA) Counts by Course
Paper-Pencil

Course	N	No. of Erasures	Mean	SD	Correlation between ERA and WTR	Number of Erasure by Percentiles						Max
						50	75	90	95	99	99.9	
9LCO	18247	12015	0.66	1.18	0.82	0	1	2	3	5	9	18
AMLC	9853	6188	0.63	1.14	0.82	0	1	2	3	5	9	15
CALG	3630	4771	1.31	1.88	0.80	1	2	4	5	8	14	23
AGEO	6724	6888	1.02	1.66	0.80	0	1	3	4	7	12	19
ALG1	9751	10084	1.03	1.58	0.81	0	2	3	4	7	12	20
GEOM	2854	2269	0.80	1.32	0.82	0	1	2	3	6	9	16
BIOL	8668	6901	0.80	1.50	0.82	0	1	2	4	7	13	27
PHSC	4718	5171	1.10	1.79	0.84	0	2	3	5	8	14	20
HIST	8492	7730	0.91	1.61	0.81	0	1	3	4	7	14	19
ECON	1763	1611	0.91	1.71	0.87	0	1	3	4	7	16	25

Table 2. State Summary Statistics for All Types of Erasure (ERA) Counts by Course
Online

Course	N	No. of Erasures	Mean	SD	Correlation between ERA and WTR	Number of Erasure by Percentiles						Max
						50	75	90	95	99	99.9	
9LCO	98287	471488	4.80	3.40	0.82	4	7	9	11	16	22	35
AMLC	87805	392559	4.47	3.30	0.83	4	6	9	11	15	21	33
CALG	22588	187173	8.29	4.79	0.73	8	11	14	17	23	33	51
AGEO	73164	540786	7.39	4.48	0.73	7	10	13	16	21	30	50
ALG1	86577	682323	7.88	4.66	0.73	7	10	14	17	22	32	57
GEOM	26546	168315	6.34	4.05	0.73	6	8	12	14	19	27	49
BIOL	99967	908488	9.09	5.77	0.79	8	12	17	20	27	39	62
PHSC	79290	705721	8.90	5.24	0.78	8	12	16	19	25	36	67
HIST	88526	800338	9.04	5.85	0.80	8	12	17	20	27	39	64
ECON	58332	554914	9.51	6.18	0.82	8	13	17	21	29	42	63

Table 3. State Summary Statistics for Wrong-to-Right (WTR) Erasure Counts by Course Paper-Pencil

Course	N	No. of Erasures	Mean	SD	Correlation between ERA and WTR	Number of Erasure by Percentiles						Max
						50	75	90	95	99	99.9	
9LCO	18247	6437	0.35	0.75	0.82	0	0	1	2	3	5	13
AMLC	9853	3433	0.35	0.72	0.82	0	0	1	2	3	5	7
CALG	3630	2145	0.59	1.02	0.80	0	1	2	3	4	7	10
AGEO	6724	3121	0.46	0.91	0.80	0	1	2	2	4	6	15
ALG1	9751	4864	0.50	0.91	0.81	0	1	2	2	4	6	12
GEOM	2854	1236	0.43	0.83	0.82	0	1	1	2	4	5	7
BIOL	8668	3213	0.37	0.81	0.82	0	0	1	2	4	6	11
PHSC	4718	2446	0.52	1.00	0.84	0	1	2	2	4	9	12
HIST	8492	3546	0.42	0.88	0.81	0	1	1	2	4	7	9
ECON	1763	786	0.45	1.03	0.87	0	1	1	2	4	10	16

Table 4. State Summary Statistics for Wrong-to-Right (WTR) Erasure Counts by Course Online

Course	N	No. of Erasures	Mean	SD	Correlation between ERA and WTR	Number of Erasure by Percentiles						Max
						50	75	90	95	99	99.9	
9LCO	98287	249202	2.54	2.23	0.82	2	4	5	7	10	15	23
AMLC	87805	218957	2.49	2.21	0.83	2	4	5	7	10	14	26
CALG	22588	83385	3.69	2.69	0.73	3	5	7	9	12	17	26
AGEO	73164	231271	3.16	2.46	0.73	3	4	6	8	11	15	23
ALG1	86577	314247	3.63	2.72	0.73	3	5	7	9	12	18	49
GEOM	26546	76920	2.90	2.33	0.73	2	4	6	7	10	15	25
BIOL	99967	433324	4.33	3.38	0.79	4	6	9	11	15	23	46
PHSC	79290	322196	4.06	2.98	0.78	4	6	8	10	13	20	48
HIST	88526	368164	4.16	3.26	0.80	3	6	8	10	15	23	52
ECON	58332	265229	4.55	3.52	0.82	4	6	9	11	16	25	52

Table 5. Number of Schools Flagged for Erasure and WTR Erasure and WTR Analysis Paper-Pencil

Course	Total Number of Schools	All Erasure Analyses		Wrong-to-Right Erasure Analyses	
		Number of Schools Flagged	Percent of Schools Flagged	Number of Schools Flagged	Percent of Schools Flagged
9LCO	176	6	3.41	1	0.57
AMLC	141	4	2.84	0	0.00
CALG	82	1	1.22	2	2.44
AGEO	99	2	2.02	2	2.02
ALG1	110	3	2.73	2	1.82
GEOM	44	2	4.55	0	0.00
BIOL	149	2	1.34	1	0.67
PHSC	132	2	1.52	1	0.76
HIST	140	2	1.43	3	2.14
ECON	85	0	0.00	1	1.18

Table 6. Number of Schools Flagged for Erasure and WTR Erasure and WTR Analysis Online

Course	Total Number of Schools	All Erasure Analyses		Wrong-to-Right Erasure Analyses	
		Number of Schools Flagged	Percent of Schools Flagged	Number of Schools Flagged	Percent of Schools Flagged
9LCO	533	11	2.06	6	1.13
AMLC	444	4	0.90	3	0.68
CALG	343	3	0.87	7	2.04
AGEO	383	13	3.39	19	4.96
ALG1	598	5	0.84	19	3.18
GEOM	169	2	1.18	3	1.78
BIOL	475	7	1.47	10	2.11
PHSC	657	10	1.52	11	1.67
HIST	449	7	1.56	6	1.34
ECON	434	5	1.15	2	0.46

Table 7. Number of Schools Flagged (WTR) in any Course for Milestones EOC

Course	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	293	11	3.75	282	96.25
Online Tests	825	61	7.39	764	92.61

8

Appendix E: 2016 DRC EOG Unusual Response Executive Report

Modified Jacob and Levitt Analyses

Submitted by DRC

January 2017

With the high-stakes nature of large-scale assessments such as the Milestones End of Grade (EOG), there are times when students' responses, and hence their scores, may not be a true representation of their own abilities. Various activities may take place, such as a student copying from another student's paper, students receiving inappropriate assistance before or during testing, or students' responses altered after testing. To maintain the integrity of the Milestones EOG and the validity of the results, it is important that any such instances be discovered.

The present study investigated student responses on the English Language/Arts and Mathematics assessments of the 2016 Spring Milestones EOG using a modified application of the analysis described by Jacob and Levitt (2003). This method is designed to identify schools with both large score fluctuations across years and unexpected patterns in student answers.

Method

This method included a combination of two indices: (1) unexpected test score fluctuations across years using a cohort of students and (2) unexpected patterns in student answers. The first indicator ranked each school's average test score gains relative to other schools' gains for a particular grade and subject. The second index ranked schools regarding unexpected patterns in student answers. The student answer pattern analyses were examined in four ways. Schools were ranked on four measures that were combined to provide an overall index of unexpected patterns in student answers. The analyses identified the:

- most unlikely block of identical answers,
- degree of correlation in student answers across the test,
- degree of variance in the correlation of responses across items, and
- extent to which student responses were congruent with respect to item difficulty and student ability.

It was possible for a school to experience a large increase in tests scores due to, for example, the introduction of a new curriculum or after-school program. It was also possible for unexpected answer patterns to appear without inappropriate behavior having occurred. For these reasons, a school had to be in the 95th percentile on both indices to be flagged. Having to be within the 95th percentile on both indices, in this context, was a way to limit the number of schools being identified due to Type I error. In this case, a Type I error would be incorrectly identifying a school for suspicious behavior. By ranking the schools on both indices Type I error is made smaller than if using only one index and the schools are protected from

being falsely identified.

INDEX 1

The first indicator ranked each school/administrator's average test score gains relative to other schools'/administrators' gains for a particular grade and subject. The mathematical form of this index is Index 1 = rank gain, where rank gain[1] is the percentile rank for average test score gains for all students in each cohort from previous year to current year as ordered by the probability of obtaining a given change or more extreme in deviations from the mean across years assuming the distribution of the test score change follows the *t* distribution. Cohorts that yield values in the top 95th percentile of this index are identified as having unusual test score fluctuations.

INDEX 2

The second index ranked schools regarding unexpected patterns in student answers. The student answer pattern analyses were examined in four ways. Schools' rankings on the four measures were combined to provide an overall index of unexpected patterns in student answers. The analyses identified the following measures:

1. most unlikely block of identical answers,
2. highly correlated answers across the test,
3. degree of variance in the correlation of responses across items, and
4. cases in which students miss easy items while answering difficult items correctly.

Measure 1 identifies the most unlikely block of identical answers given by students on consecutive items using a multinomial logit model. The likelihood of each student choosing each possible answer on every item is calculated based on the student's current year's test responses and previous year's test scores. All combinations of students and consecutive items are compared to find the block of identical answers that were least likely to have arisen by chance.

First, a multinomial logit model is used to calculate every student's likelihood on each item:

$$1.\quad P(Y_{isc} = k) = \frac{e^{\beta_k x_s}}{\sum_{j=1}^{J} e^{\beta_j x_s}}, \qquad k = 1, \ldots, J$$

where s is the student, c is the cohort, k is the selected answer option, J is the total number of options, and x is the vector of previous year's test scores.

[1] The grade equivalence (GE) was used by Jacob and Levitt for gain scores over years.

Second, the likelihood of a student's answer for item *i* is found by selecting the appropriate value from Equation 1:

2. $P_{isck} = \frac{e^{\beta_k x_s}}{\sum_{j=1}^{J} e^{\beta_j x_s}}$, where *k* is the response actually chosen by student *s* on item *i*.

Third, identify strings of items, *m* to *n*, for which the cohort gave identical responses; then the likelihood of this string for student *s* is the product of the item likelihoods from Equation 2:

3. $P_{sc}^{mn} = \prod_{i=m}^{n} P_{isck}$.

Fourth, the product across all students in the cohort who had identical responses in the string is

4. $\tilde{P}_{sc}^{mn} = \prod_{s\in\omega} P_{sc}^{mn}$, where □□is the group of students who have identical responses to items *m* to *n.* The calculations are repeated for all strings of five consecutive operational items.

Finally, the minimum value of this measure for each cohort is recorded as Measure 1.

5. *Measure 1* = $min_s\left(\tilde{P}_{sc}^{mn}\right)$

The smallest values are associated with more improbable answer strings within a cohort.

Measure 2 examines the degree of correlation in student responses across the test, particularly for unexpected answers. It was based on the assumption that teachers who cheated will have students with highly correlated answers. Measure 2 is the average of the item residual values. Higher values indicate cohorts with highly correlated answers.

6. $e_{jisc} = \begin{matrix} 0 - P_{isck} & if\ j \neq k \\ 1 - P_{isck} & if\ j = k \end{matrix}$.

where e_{ijsc} is a residual for item *i* for *j*th item answer option for cohort student sc, and P_{isck} is the probability of the cohort student, *sc*, select the answer option *k* for item *i*. Then residual for each option are summed across students within the cohort:

7. $e_{jic} = \sum_s e_{jisc}$.

This sum of residual, e_{ijsc} should be approximately zero if there is no within-school/administrator correlation in the way students responded to item *i*, response *j*. The residual for all possible responses were summed for

each item within school/administrator. Then, the sum of squared residuals is divided by the squared number of students to normalize for cohort size. This is analogous to average test residual for cohort.

8. $v_{ic} = \frac{\sum_j e_{jic}^2}{n^2}$.

Measure 2 is average value of v_{ic} across all items for cohort, and a measure of correlation across test.

9. *Measure 2* = $\bar{v_c} = \frac{\sum_i v_{ic}}{ni}$.

Measure 3 calculates the variance of test residual for cohort, v_{ic}. If a teacher cheated by changing or providing answers for multiple students on selected questions, the within-cohort residual deviation on those particular questions will be extremely high, while the within-cohort residual deviation on other questions is likely to be normal. Thus, a large degree of variance in the residual of responses across items would occur. The variance is calculated as follows:

10. *Measure 3* = $\sigma_{v_c} = \frac{\sum_i (v_{ic} - \bar{v}_c)^2}{ni}$, where *ni* is the number of items on the exam.

Measure 4 compares the answers of students within a cohort to the answers from other students with same total scores in the sample. It detects students who missed easy items while answering difficult items correctly.

For each student in cohort, deviation of item response from all students with aggregate score A is computed as follows:

11. $z_{sc} = \sum_i (q_{isc} - \bar{q}_i^A)^2$,

where superscript A indicate all students with aggregate total score A; q_{isc} equals one if the student answers item i correctly and zero otherwise; and $\bar{q}_i^A$ is the proportion of A students answering item i correctly. Squared deviation is summed for items, making test level deviation for each student. Measure 4 is average value of test level deviation from students with total score A for each cohort.

12. *Measure 4* = $\frac{\sum_S (Z_{sc} - \bar{z}^A)}{n_{sc}}$

High values of this index indicate the answer from a large number of students in the cohort deviated from students with same total scores in other cohorts.

With all four measures computed, the schools/groups are ranked on each of the four measures. For this, quantile regression is used to remove the effect of sample size inherent in the indexes used by Jacob and Levitt (2003). While *least squares* regression minimizes the sum of squared deviations from the regression line and passes through the mean, quantile regression on the median, for example, minimizes the sum of absolute values of the deviations from the line which is the median (Koenker, 2005). The result is that exactly 50% of the data points will be above the line and 50% below. It is also true that 50% of the points are expected to be above the line for any value of the independent variable. In other words, the quantile regression line is the median of the dependent variable conditional on the independent variable.

In the quantile analysis, the Modified Jacob and Levitt measures were the dependent variables and school enrollment was the independent variable. Quantile regression can be generalized to any percentile, typically denoted as □□between 0 and 99. The preceding discussion used the median (□ = 0.5). By iterating on □□using the *R* package *quantreg* (R Development Core Team, 2003; Koenker, 2011), it is possible to determine the percentile rank for any value of the measure conditional on school enrollment. The percentile ranks for each cohort, on each measure, are then combined to form the Index 2 as follows:

$$13.\ Index\ 2 = Measure1\_rank^2 + Measure2\_rank^2 + Measure3\_rank^2 + Measure4\_rank^2.$$

COMBINING INDEX 1 and INDEX 2

It was possible for a school to experience a large increase in tests scores due to, for example, the introduction of a new curriculum or after-school program. It was also possible for unexpected answer patterns to appear without inappropriate behavior having occurred. For these reasons, a school had to be in the 95th percentile on both indexes to be flagged. Having to be within the 95th percentile on both indexes, in this context, was a way to limit the schools being identified due to Type I error.

Results

Unlike the other forensic methods, for a school to be identified by this analysis it had to have a minimum sample size of eight and be in the 95th percentile for both unexpected score changes and unexpected patterns in student response indices. It is expected that fewer schools would be identified with this method since it is designed to detect only extreme cases of potential misconduct.

As can be seen in Table 1, the number of schools flagged in ELA ranged from 4 schools to 12 and from 7 schools to 23 schools in Mathematics. Grade 3 was excluded from this method since there was no data from an immediately preceding year to use in the analysis. Note that Jacob and Levitt analyses have not been performed previously within the Georgia Milestones. In subsequent administrations, we can evaluate whether the number of schools flagged within the Jacob and Levitt analyses is trending upwards or downwards.

Table 1. Count of Schools Flagged within 2016 Jacob and Levitt Analyses

Grade	English Language Arts				Mathematics			
	Form A		Form B		Form A		Form B	
	Count	%	Count	%	Count	%	Count	%
3	-	-	-	-	-	-	-	-
4	10	0.81	5	0.41	18	1.47	23	1.88
5	12	0.98	7	0.58	19	1.55	13	1.07
6	8	1.43	4	0.72	8	1.44	9	1.62
7	4	0.74	4	0.74	7	1.29	7	1.31
8	4	0.74	4	0.74	11	2.04	10	1.87
TOTAL	38	0.94	50	0.64	39	1.56	51	1.55

Discussion

The goal of psychometric forensic analysis was to screen for test results that may have been spurious because valid inferences cannot be made from such test scores. The Jacob and Levitt analyses reported here are just one set of forensic analyses implemented within the Georgia Milestones and should not be considered in isolation.

Note that the schools were flagged within the Jacob and Levitt analyses based on statistical evidence alone. If flagged, that does not necessarily mean that the schools engaged in inappropriate testing activity. However, the statistical evidence does suggest that something aberrant or unusual occurred and, barring a simple explanation, warrants further exploration.

All forensic results should be used with caution, and data for schools and grades within schools and their results may serve as good starting points for the evaluation of potential testing irregularities. The 2016 Jacob and Levitt analyses represent an important step in helping to maintain the integrity of future administrations of the Milestones EOG.

References

Jacob, B. & Levitt, S (2003). Rotten apples: An Investigation of the prevalence and predictors of teacher cheating, *The Quarterly Journal of Economics*, 118 (3), 843-877.

Koenker, R. (2005). *Quantile regression*. Cambridge, UK: Cambridge University Press.

Koenker, R. (2011). *Quantile regression in R: A vignette*. Retrieved November 1, 2011, from http://www.econ.uiuc.edu/~roger/research/rq/vig.pdf.

R Development Core Team. (2008). *R: A language and environment for statistical computing*. Retrieved from the Vienna University, Institute for Statistics and Mathematics Web site: http://www.R-project.org.