DATA FORENSICS

TECHNICAL REPORT

for the
Pennsylvania
System of School Assessment

2009 Reading and Mathematics
Grades 3, 4, 5, 6, 7, 8, and 11

Provided by
Data Recognition Corporation

July 2009

Table of Contents

List of Tables

List of Figures

Chapter One: An Introduction to Data Forensics

Under the high-stakes testing environment of NCLB, many inferences (e.g., decisions about the quality of schools or whether a student can advance to the next grade) are made based on test results. The integrity of test results, therefore, is paramount. Test results must be earned fairly to allow for proper and valid inferences (i.e., decisions) from those test results. Thus, it is important to look at the test results for potential indications of irregularities at the student or school level before results are sanctioned and decisions made based on them.

Statistical indices at the student level can indicate whether a student may be cheating on the test. At the student level, one can examine the student response pattern to see if it is consistent with that predicted by the underlying mathematical model used to create the test. One can also examine the pattern of erasures exhibited by the student. Alone, each index may suggest an irregular testing behavior that should be investigated. Together, they provide more support to the notion that the test results were suspiciously earned. Though the response pattern and erasure analyses may implicate testing irregularity by the student, they may also implicate a school official who may have committed the act of erasing the student answer document.

At the school level, for each grade, it is important to examine whether the scale scores and percentages of proficient and advanced students change dramatically from one year to the next. Drastic changes in test participation rates may also be indications of suspicious activity. Together, these indices can be used to screen for potential test results that may have been earned unfairly. This is the goal of psychometric data forensics. Using various statistical indices, potential testing irregularities that may have been committed at the school or student level can be identified for further investigation.

PROPER USES OF DATA FORENSICS RESULTS

An important policy question is how one should proceed if there are statistical indications of irregularities at the student or school level. Caution must be observed when using results from data forensics. Unless the statistics can be supported by more direct evidence, such as proctor reports, punitive action is probably ill advised. Ideally, knowledge that testing irregularity analyses are conducted will be sufficient deterrent, but as the stakes increase, risk taking might as well. The use of the multiple indicators is especially important in this respect. It is also important for the forensics analyses to operate as deterrents rather than detectors.

In addition to considering how to react to specific information of testing irregularities, care should be taken to decide on the most effective communications strategies, both before and after the fact. Before testing, the communications should emphasize a positive message that some very sophisticated techniques are being employed to ensure that the assessment treats everyone fairly and equally. While the communication should be positive, the underlying message is that attempts to subvert the assessment will be detected and acted upon. The primary value of communications after the fact should be to strengthen the deterrent for future administrations. At the same time, care must be taken in these messages to protect the rights and privacy of any individuals involved.

IMPROPER USES OF DATA FORENSICS RESULTS

Student scores or school AYP status should not be revoked or invalidated based on data forensic analysis results alone. These results are based on statistical properties. Schools or students with

highly improbable results are identified. These flagged or identified schools or students may have earned their scores unfairly. However, they also may have earned them fairly. As stated earlier, the combination of many indices may give strong evidence for a testing irregularity, but they do not provide definitive evidence of such behavior.

Chapter Two: Method

Data forensic analyses of the PSSA 2009 data were analyzed using DRC's data forensics software system PSI (Psychometric Scenes Investigator). The PSI Data Forensics system is comprised of the following components: Erasure Analyzer, Residuals Analyzer, AYP I Data Analyzer, and AYP II Data Analyzer. Brief descriptions of each component follow.

The **Erasure Analyzer** determines if the numbers of erasures from wrong-to-right are improbable. The number of erasures is counted during scanning. Schools or students with an improbable number of wrong-to-right erasures are identified. The aberrance thresholds are determined using statewide data.

The **Residuals Analyzer** determines if the difference between the observed and expected outcomes (i.e., residual) for a student is improbable. The residuals are summed across operational items. Students with improbable residuals are identified. These analyses are conducted at the student level.

The **AYP I Data Analyzer** determines if the changes in scores and AYP subgroup participation have improbably changed *across* years. Schools are flagged for improbable changes in scale scores, the percentage of students who are proficient or above, and subgroup participation. Schools with a high number of wrong-to-right erasures are also identified. These analyses are conducted at the grade level for each school.

The **AYP II Data Analyzer** determines if the changes in scores and AYP subgroup participation have improbably changed *within* year. Schools are flagged for improbable changes in scale scores, the percentage of students who are proficient or above, and subgroup participation. Aberrant schools are flagged. Similar to the AYP I Data Analyzer, these analyses are conducted at the grade level for each school level.

DATA

The following data files were used to conduct the data forensic analyses:

1. 2008 Final PDE Student File – this file was used in the AYP I Data Analyzer.
2. 2009 Pre-Attribution PDE Student File – this file was used in the AYP II Data Analyzer.
3. 2009 Post-Attribution PDE Student File – this file was used in the AYP I Data Analyzer and the AYP II Data Analyzer.
4. 2009 Reading and Mathematics Erasure Analysis File – this file was used in the Erasure Analyzer.

THE THREAT SCALE

The common thread across all the different components of the PSI system is to find *aberrant* behavior that may have been carried out by a student or a school administrator. Aberrant behavior is defined as something this is not typical; that is, it is something that is abnormal, unusual, or uncharacteristic.

In the realm of test results, what constitutes aberrant behavior? How can aberrance be operationally defined? In DRC's PSI system, aberrance is defined as behavior that poses a *threat*

to the validity of the score. That is, aberrant scores are so uncharacteristic of other scores that they do not allow for proper and valid inferences to be made from the scores.

For the PSSA data forensics project, an aberrant behavior is defined as one that occurs in less than 1 in 10,000 chances. So, any behavior that would have a chance occurrence of less than 0.0001 is considered aberrant.

Small probabilities are difficult to process and display. To facilitate easy understanding and interpretation of very small numbers (i.e., very small probabilities), the *Threat Score* was created. The Threat Score (TS) is computed using the following equation:

(1). $$TS = \left| 1.086 \ln\left(\frac{p}{q}\right) \right|$$, where

p is the probability of the occurrence of the behavior, and
q is 1-*p*.

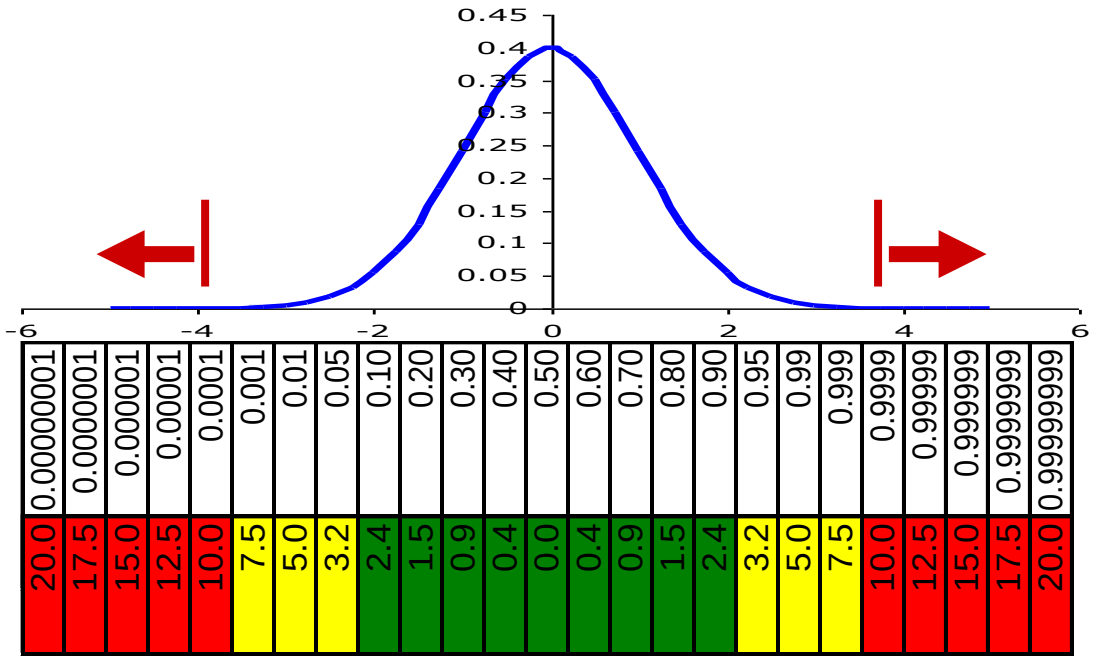
The natural log of *p/q* is taken to make the scale symmetric around small and large probabilities. The constant 1.086 is used to make the probability of 0.0001 (the aberrant criterion) equal to a threat score of 10. Thus, for ease of use and interpretation, any behavior that results in a threat score of 10 or greater is considered aberrant and a threat to the validity of the score.

The table below shows the Threat Score for selected probabilities from 0.00000001 to 0.99999999. Note that the very extreme scores (very small probabilities) have a threat score of 10 or higher and are colored red. Threat scores from 3.2 to 7.5, ranging from p-values of 0.05 to 0.0001 are colored yellow. Threat scores with probabilities ranging from 0.10 to 0.90 are colored green. Figure 1 shows the properties of the threat score scale in relation to the normal distribution (i.e., normal curve).

Table 1. Selected Probabilities for Threat Scores from 0 – 20.

p	q	ln(p/q)	Threat Score	1 in …
0.00000001	0.99999999	-18.42	20.0	100,000,000
0.0000001	0.9999999	-16.12	17.5	10,000,000
0.000001	0.999999	-13.82	15.0	1,000,000
0.00001	0.99999	-11.51	12.5	100,000
0.0001	0.9999	-9.21	10.0	10,000
0.001	0.999	-6.91	7.5	1,000
0.01	0.99	-4.60	5.0	100
0.05	0.95	-2.94	3.2	
0.10	0.90	-2.20	2.4	
0.20	0.80	-1.39	1.5	
0.30	0.70	-0.85	0.9	
0.40	0.60	-0.41	0.4	
0.50	0.50	0.00	0.0	
0.60	0.40	0.41	0.4	
0.70	0.30	0.85	0.9	
0.80	0.20	1.39	1.5	
0.90	0.10	2.20	2.4	
0.95	0.05	2.94	3.2	
0.99	0.01	4.60	5.0	100
0.999	0.001	6.91	7.5	1,000
0.9999	0.0001	9.21	10.0	10,000
0.99999	0.00001	11.51	12.5	100,000
0.999999	0.000001	13.82	15.0	1,000,000
0.9999999	0.0000001	16.12	17.5	10,000,000
0.99999999	0.00000001	18.42	20.0	100,000,000

Figure 1. Threat Score Scale.



ERASURE ANALYZER

The objective of this analysis is to find potential testing irregularities that may have been perpetrated by a student or someone with access to a student's answer document. Students with an aberrant number of wrong-to-right (WR) erasures are identified. To determine if the student's number of erasures is aberrant, the student's erasures are compared to the state mean and standard deviation. The following statistics are calculated:

(2). $Mean = \frac{\sum X_i}{N}$, where

X_i = number of W-R erasures for student *i*, and
N = total number of students in the state.

(3). $SD = \sqrt{\frac{\sum_{i=1}^{k} (X_i - Mean)^2}{N-1}}$, where

Mean = mean number of erasures in the state.

(4). $z = \frac{x_i - Mean}{SD}$,

(5). p = the probability of the z statistic, which is converted into the Threat Score using Equation (1).

Criteria for computing statistics and for flagging students:

- Include all students with a valid test score
- Include only operational items in count of WR erasures
- Flag student if $TS \geq 50$

Note that we are using a threat score of 50 to flag students in the erasure analysis. Using a flagging criterion of 10 resulted in too many students being flagged. In this first year of data forensics for the PSSA, we wanted to keep the number of flagged students to a manageable number.

Also, note that we are also flagging a school for erasure analysis. That is, we are flagging schools with an aberrant (i.e., high number) mean number of erasures that went from wrong to right. To flag a school for erasure analysis, the following statistics are computed:

(6). $Mean_{School} = \frac{\sum X_i}{n}$, where

X_i = number of WR erasures for student *i* in the school, and
n = total number of students in the school.

(7). $SD_{School} = \sqrt{\frac{\sum (X_i - Mean_{School})^2}{n-1}}$, where

(8). $t = \frac{Mean_{School} - Mean}{\sqrt{\frac{SD_{School}}{n-1}}}$, where degrees of freedom (df) is n-1

(9). p = the probability of the t statistic, which is converted into the Threat Score using Equation (1).

In combination, school and student erasure analysis flags may be good indicators of potential testing irregularities. Thus, if a school is flagged for erasure analysis, there are likely many students who were also flagged. It is these schools and students that may warrant further investigation.

For students who are flagged for having an aberrant number of WR erasures, the specific type of erasure is shown in the student report. For each item, the type of erasure, if any, is indicated, with WR indicating a wrong-to-right erasure, WW indicating a wrong-to-wrong erasure, and RW indicating a right-to-wrong erasure. The number of WR, RW, and total erasures are shown in the report.

Criteria for computing statistics and for flagging schools:

- Include all students with a valid test score
- Include only operational items in count of W-R erasures
- Include only students with full academic year status
- Flag school if $TS \geq 10$

RESIDUAL ANALYZER

Similar to the Erasure Analyzer, the objective of this analysis is to find potential testing irregularities that may have been perpetrated by a student or someone with access to a student's answer document. The heart of the system is Rasch residual analysis (Mead, 1976, 1980; Smith, 1982, 1986, 2004). These techniques have been used successfully by a number of high stakes certification and licensing agencies. In these assessments, they have detected examinee copying, distribution of secure items within schools, obsolete items, and processing errors such as mismatching candidate records and incorrect form assignment.

The defining characteristic of Rasch measurement is *specific objectivity*. For the individual, this means that the probability of success on any item depends only on the person's ability and the item's difficulty. After extracting the information associated with the sufficient statistic for estimating ability, the remainders can be used to control the model by assessing the degree to which specific objectivity is obtained. Inspecting the residuals can provide valuable diagnostic information.

The fundamental unit of data is the response of one person to one item. Because both the person ability and the item difficulty have been estimated from larger data sets, they can be effectively removed from the observation leaving the typical person-item residual. The following statistics are computed:

(10). $y_{vi} = x_{vi} - p_{vi}$, where $p_{vi} = \dfrac{e^{b_v - d_i}}{1 + e^{b_v - d_i}}$.

(11). $z_{vi} = \sqrt{\dfrac{1 - p_{vi}}{p_{vi}}} = \sqrt{\dfrac{e^d}{e^b}}$, for correct response, where

b is the student ability, and
d is item difficulty.

(12). $z_{vi} = -\sqrt{\dfrac{p_{vi}}{1 - p_{vi}}} = -\sqrt{\dfrac{e^b}{e^d}}$, for an incorrect response, where

b is the student ability, and
d is item difficulty.

(13). $outfit_{zstd} = (MS^{\frac{1}{3}} - 1)(\dfrac{3}{S}) + (\dfrac{S}{3})$

(14). $MS = \dfrac{\sum_{i=1}^{N} z_{vi}^2}{N}$, where N is the number of points.

(15). $S = \dfrac{\left[\sum \dfrac{1}{w_{ni}} - 4N\right]^{\frac{1}{2}}}{N}$

(16). $w = p_{vi}(1 - p_{vi})$

(17). p = the probability of the $outfit_{zstd}$ statistic based on the standard normal cumulative distribution, which is converted into the Threat Score using Equation (1).

For students who are flagged by the Residual Analyzer, the residual for each item is shown in the student report. Residuals are rounded to the nearest integer. Positive residuals indicate that the student performed better than was predicted by the Rasch model. On the other hand, negative item residuals indicate that the student performed less well than the model would suggest. High positive or low negative numbers are more indicative of misfit to the model. Item residuals ranging from -1 to +1 are typical and should not be of much concern.

Criteria for computing statistics and for flagging students:

- Include all students with a valid test score
- Include only operational items in residuals
- Include only students with full academic year status
- Flag student if $TS \geq 10$

AYP I DATA ANALYZER

The objective of these analyses is to find potential testing irregularities at the school level. Any school-level testing irregularities may render the AYP status of a school invalid. Thus, it is critical that school level test results are valid and earned fairly. Note that unlike the AYP status calculations these analyses are conducted at the grade level. Data forensic results are not aggregated at the school level. Examining the results at a finer grain level allows us to better observe the testing behaviors at the school. This component of the PSI system looks for aberrant results in:

- Scale score change from year to year
- Percent Proficient+Advanced change from year to year
- Regression residual
- Subgroup participation percentage changes from year to year

School Scale Score Change

To determine whether a school has an improbable change in scale score, the means of two independent samples is compared by conducting a t-test. The following statistics are calculated:

(18). $$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$, where

$\bar{X}_1$ is the mean score of each school of first year,

$\bar{X}_2$ is the mean score of each school of second year,

μ_1 is the mean score of state level of first year,

μ_2 is the mean score of state level of second year, and

(19). $s_p^2 = \frac{(n_1 - 1) \times s_1^2 + (n_2 - 1) \times s_2^2}{n_1 + n_2 - 2}$, where

s_1^2 is the variance of each school of first year,

s_2^2 is the variance of each school of second year,

n_1 =the number of students of each school of first year, and

n_2 =the number of students of each school of second year

(20). p = the probability of the t statistic, where df is ($n_1 + n_2 - 2$), which is converted into the Threat Score using Equation (1).

Criteria for computing statistics and for flagging schools:

- Include all students with a valid test score
- Include only students with full academic year status
- If $n_1 \geq 10$ and $n_2 \geq 10$, conduct the above calculations and flag school if $TS \geq 10$
- If $n_1 < 10$ or $n_2 < 10$, calculate change in scale score, do not conduct t-test, and set $TS = 0$.

School Percent Proficient+Advanced (P+A) Change

To determine whether a school has an improbable change in the percentage of students are who proficient or advanced, the log odds ratio is used to compare the percentage of students who are proficient or advanced in the current year to the percentage in the previous year.

For a probability π of proficient+advanced, the odds are defined to be

(21). $\Omega = \frac{\pi}{1 - \pi}$.

The ratio of the odds Ω_1 (for current year) and Ω_2 (for previous year) is

(22). $\theta = \frac{\Omega_1}{\Omega_2} = \frac{\pi_1 / 1 - \pi_1}{\pi_2 / 1 - \pi_2}$.

The odds ratio is also called the cross-product ratio, since it equals the ratio of the products $\pi_{11}\pi_{22}$ and $\pi_{12}\pi_{21}$ of cell probabilities from diagonally opposite cells. The odds ratio can equal any nonnegative number. The conditions $\Omega_1 = \Omega_2$ and $\theta = 1$ correspond to independence of the current and previous year. When $1 < \theta < \infty$, students in the current school year are more likely to have proficient+advanced scores than are students in the previous school year. For example, when $\theta = 4$, the odds of proficiency in current school year are four times the odds in previous school year.

Values of θ farther from 1.0 in a given direction represent stronger association. Different direction of values represents the same association, but in opposite direction, when one is the inverse of the other. For example, θ =0.25, the odds of proficiency in the current school year are 0.25 times the odds in the previous school year, or equivalently, the odds of proficiency in the previous school year are 1/0.25=4.0 times the odds in the current school year.

For inference it is convenient to use $\log\theta$ (Agresti, 2002). Independence corresponds to $\log\theta$ = 0. The log odds ratio is symmetric about this value. The following statistics are used to calculate the log odds ratio:

(23). $\ln(\hat{\theta}) = \ln\left[\frac{(n_{11}+0.5)(n_{22}+0.5)}{(n_{12}+0.5)(n_{21}+0.5)}\right]$, where

n_{11} = the number of students P+A in first year
n_{12} = the number of students NOT P+A in first year
n_{21} = the number of students P+A in second year
n_{22} = the number of students NOT P+A in second year

(24). $\hat{\sigma} = \left(\frac{1}{n_{11}+0.5} + \frac{1}{n_{12}+0.5} + \frac{1}{n_{21}+0.5} + \frac{1}{n_{22}+0.5}\right)^{1/2}$

(25). $z = \frac{\ln(\hat{\theta})}{\hat{\sigma}}$

(26). p = the probability of the z statistic, which is converted into the Threat Score using Equation (1).

Note that the 0.5 in Equation (23) and (24) are used to deal with problems that may occur with using log odds ratios with small sample size. It is preferred when the cell counts are very small or any zero cell counts occur. Gart and Zweifel (1967) showed that these amended estimators behave well.

Criteria for computing statistics and for flagging schools:

- Include all students with a valid test score
- Include only students with full academic year status
- If $n_{1+} \geq 10$ and $n_{2+} \geq 10$, conduct the above calculations and flag school if $TS \geq 10$
- If $n_{1+} < 10$ or $n_{2+} < 10$, calculate the percentage proficient+advanced, do not calculate the log odds ratio, and set $TS = 0$.

School Regression Residual

The mathematics and reading scale scores for a school are often correlated. Thus, it is possible to predict one score from the other. When the actual mean score for a school is hugely different from the predicted mean score, the validity of the actual (observed) score may be questionable. More specifically, when the observed score is so much greater than the predicted score, this pattern of results could suggest that teachers or school officials were engaged in inappropriate testing behavior in one subject area and not the other. To determine if the difference between the observed and predicted scores are aberrant, a regression analysis is conducted to identify schools that have large residuals between observed and predicted scores. The equations below show how the regression is done to predict the reading scores from the mathematics scores. Similar analyses are conducted to predict the mathematics scores from the reading scores.

(27). $\hat{Y}_j = \beta_0 + \beta_1 X_j$, where

$\hat{Y}_j$ is the estimated reading scale score of student *j* in the state level

X_j is the actual mathematics scale score of student *j* in the state level

β_0 is the intercept for the state level

β_1 is the slope for the state level.

(28). $\hat{\bar{y}}_k = \beta_0 + \beta_1 \bar{x}_k$, where

$\hat{\bar{y}}_k$ is the predicted mean reading scale score of school *k*

$\bar{x}_k$ is the actual mean mathematics scale score of school *k.*

(29). $t = \dfrac{\bar{y}_k - \hat{y}_k}{s_{\bar{y}}}$, where

$\hat{\bar{y}}_k$ is the predicted mean reading scale score of school *k*

$\bar{y}_k$ is the actual mean reading scale score of school *k,* and

(30). $s_{\bar{y}} = \sqrt{\dfrac{\sum (y - \hat{\bar{y}})^2}{n-1}}$, where

n is the number of students in the school.

(31). *p* = the probability of the *t* statistic, where *df* is $(n-1)$, which is converted into the Threat Score using Equation (1). Negative values of *t* are converted to a T*hreat Score* of 0, since we only want to flag a school when the observed score is much greater than the predicted score.

Criteria for computing statistics and for flagging schools:

- Include all students with a valid test score
- Include only students with full academic year status
- If $n_1 \geq 10$ and $n_2 \geq 10$, conduct the above calculations and flag school if $TS \geq 10$
- If $n_1 < 10$ or $n_2 < 10$, calculate predicted s core, do not conduct t-test, and set $TS = 0$.

School Subgroup Participation Change

Subgroup participation rates are expected to be consistent from one year to the next. Huge changes in participation may suggest irregular testing activities or coding of student demographics. To determine whether subgroup participations rates are aberrant, the log odds ratio is used to compare the percentages of students in the subgroup in the current year to the percentages in the previous year. The analyses are conducted for the following subgroups: ELL, IEP, ED, AI/AN, ASIAN, BLACK, HISP, WHITE, and MULT. The following statistics are computed:

(32). $\ln(\hat{\theta}) = \ln\left[\dfrac{(n_{11}+0.5)(n_{22}+0.5)}{(n_{12}+0.5)(n_{21}+0.5)}\right]$, where

n_{11} = the number of students subgroup in first year,
n_{12} = the number of students NOT in subgroup in first year,

n_{21} = the number of students subgroup in second year,
n_{22} = the number of students NOT in subgroup in second year.

(33). $$\hat{\sigma} = \left(\frac{1}{n_{11}+0.5} + \frac{1}{n_{12}+0.5} + \frac{1}{n_{21}+0.5} + \frac{1}{n_{22}+0.5} \right)^{1/2}$$

(34). $$z = \frac{\ln(\hat{\theta})}{\hat{\sigma}}$$

(35). p = the probability of the z statistic, which is converted into the Threat Score using Equation (1).

Criteria for computing statistics and for flagging:

- Include all students with a valid test score
- Include only students with full academic year status
- If $n_{1+} \geq 10$ and $n_{2+} \geq 10$, conduct the above calculations and flag school if $TS \geq 10$
- If $n_{1+} < 10$ or $n_{2+} < 10$, calculate the change percentage, do not calculate the log odds ratio, and set $TS = 0$.

AYP II DATA ANALYZER

The objective of these analyses is to find potential irregularities with the testing results that may occur during the attribution process. Thus, these analyses look for aberrant results within year by comparing the pre-attribution and the post-attribution data files. Again, these results are provided at the grade level for each school. That is, data forensics results are not aggregated at the school level. This component of the PSI system looks for aberrant results in the following areas:

- Percent of students tested who do not have full academic year status
- Scale score changes
- Percent Proficient+Advanced changes
- Subgroup participation changes

Percent of Students Tested Who are Full Academic Year

Of all the students who tested, most are expected to have full academic year (FAY) status. Only results from FAY students are counted towards adequate yearly progress (AYP) status. Thus, when a very low percentage of the students tested are FAY, it may indicate an aberrant testing pattern that should be further investigated.

To determine whether a school has an aberrant percentage of testers who are FAY, the school's percentage of FAY testers is compared to the state percentage. The following statistics are calculated:

(36). $x_i = \frac{n_{NFAY}}{n_{TOTAL}} * 100$, where x_i is the percentage of FAY testers for a school,

where n is the number of students in the school.

(37). $Mean = \frac{\sum x_i}{N}$, is the mean percentage of FAY testers for the state.

(38). $SD = \frac{\sum_{i=1}^{k} (x_i - Mean)^2}{N-1}$, is the standard deviation of FAY testers for the state,

where N is the number of schools in the state.

(39). $t = \frac{x_i - Mean}{\frac{SD}{\sqrt{n}}}$.

(40). p = the probability of the t statistic, which is converted into the Threat Score using Equation (1).

For this check, there is no absolute value sign in the Threat Score Equation (1). We only flag a school when the school percentage is much lower than the state mean. We do not care if it is greater than the state mean (in this case, it is good for the school to be close to 100% FAY, meaning all testers are FAY students).

Criteria for computing statistics and for flagging schools:

- Include all students with a valid test score
- Flag school if $TS \geq 10$

School Scale Score Change

To determine whether a school has an improbable change in scale score after the attribution process, the means of two independent samples is compared by conducting a t-test. The following statistics are calculated:

(41). $t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$, where

$\bar{X}_1$ is the mean score of each school pre-attribution,

$\bar{X}_2$ is the mean score of each school of post-attribution,

μ_1 is the mean score of state level of pre-attribution,

μ_2 is the mean score of state level of post-attribution, and

(42). $s_p^2 = \frac{(n_1 - 1) \times s_1^2 + (n_2 - 1) \times s_2^2}{n_1 + n_2 - 2}$, where

s_1^2 is the variance of each school pre-attribution,

s_2^2 is the variance of each school post-attribution,

n_1 =the number of students of each school pre-attribution,

n_2 =the number of students of each school post-attribution

(43). p = the probability of the t statistic, with $df = (n_1 + n_2 - 2)$, which is converted into the Threat Score using Equation (1).

Criteria for computing statistics and for flagging schools:

- Include all students with a valid test score
- Include only students with full academic year status
- If $n_1 \geq 10$ and $n_2 \geq 10$, conduct the above calculations and flag school if $TS \geq 10$
- If $n_1 < 10$ or $n_2 < 10$, calculate change in scale score, do not conduct t-test, and set $TS = 0$.

School Percent Proficient+Advanced (P+A) Change

To determine whether a school has an improbable change in the percentage of students are who proficient or advanced after the attribution process, the log odds ratio is used to compare the percentage of students who are proficient or advanced in the pre-attribution file to the percentage in the post-attribution file. The following statistics are used to calculate the log odds ratio:

(44). $\ln(\hat{\theta}) = \ln\left[\dfrac{(n_{11}+0.5)(n_{22}+0.5)}{(n_{12}+0.5)(n_{21}+0.5)}\right]$, where

n_{11} = the number of students P+A pre-attribution
n_{12} = the number of students NOT P+A pre-attribution
n_{21} = the number of students P+A post-attribution
n_{22} = the number of students NOT P+A post-attribution

(45). $\hat{\sigma} = \left(\dfrac{1}{n_{11}+0.5} + \dfrac{1}{n_{12}+0.5} + \dfrac{1}{n_{21}+0.5} + \dfrac{1}{n_{22}+0.5}\right)^{1/2}$.

(46). $z = \dfrac{\ln(\hat{\theta})}{\hat{\sigma}}$.

(47). p = the probability of the z statistic, which is converted into the Threat Score using Equation (1).

Criteria for computing statistics and for flagging:

- Include all students with a valid test score
- Include only students with full academic year status
- If $n_{1+} \geq 10$ and $n_{2+} \geq 10$, conduct the above calculations and flag school if $TS \geq 10$
- If $n_{1+} < 10$ or $n_{2+} < 10$, calculate the percentage proficient+advanced, do not calculate the log odds ratio, and set $TS = 0$.

School Subgroup Participation Change

For AYP calculations, schools[1] are only accountable for subgroups with 40 or more students. When the number of students in a subgroup meets this criterion in the pre-attribution file but does not in the post-attribution file, this change may indicate an unsanctioned or unwarranted change to a student's demographic profile or user error in the attribution process. In either case, it may be worth investigating these instances. To check if the number of students tested changes from being above to being below the critical value of 40, the following statistics are computed for the following subgroups (ELL, IEP, ED, AI/AN, ASIAN, BLACK, HISP, WHITE, MULT):

(48). N_{pre}, the number of students in the pre-attribution file.

(49). N_{post} , the number of students in the post-attribution file.

Criteria for computing statistics and for flagging schools:

- Include all students with a valid test score
- Include only students with full academic year status
- Flag the school if $N_{pre} \geq 40$ and $N_{post} < 40$.

[1] Note that since data forensic results are provided at the grade level, these analyses at the grade level may not be as informative. In hindsight, it may have been better to aggregate these results at the school level.

Chapter Three: Results

The results of the data forensic analyses are provided as HTML reports. Open the file "Enter Data Forensics Site.html" in the folder "PSSA 2009 – Data Forensics" to start viewing the results. It is important to keep the files and folder structure as is (i.e., do not move or rename files and folders). Keeping the file and folder structure as it is will allow the user to navigate to a specific district, school, or student report via the left task bar. There should be no need to move any files around.

ORGANIZATION OF RESULTS IN HTML FORMAT

For each school, results are provided *by grade level*. In contrast, AYP status results are provided at the school level. For data forensics, it is important to see results at a finer grain level. Thus, results are provided at the student level when applicable (for residual and erasure analyses) and at the grade level. The State Report gives a summary of the state result for that grade. Separate state reports are provided for the AYP I and AYP II analyses. When viewing the AYP I report, click on *View: AYP II* to see the AYP II results. Similarly, when viewing the AYP II report, one can navigate back to the AYP I report by clicking on *View: AYP I*.

From the state report, one can navigate to a specific grade report, or to an index of districts. From the district index, one can select from the list of districts, which are listed in alphabetical order. The district report gives a list of the schools in the district that was flagged by one or more of the indicators. One can navigate to a specific school by directly clicking on the school name.

The school report shows detailed results from the data forensic analyses. If a threat score for an analysis is greater than the criterion of 10, there is an "X" under the "Flag" column. A red highlighted row also indicates that a school was flagged for that analysis. When there are students in the school that were flagged by the residual or erasure analyses, a list of flagged students is provided. Clicking on the student ID will bring the detailed report for the student.

Summary of Flagged Schools

For each of the analyses, a summary of the schools that were flagged by each indicator can be viewed by clicking on "Flag Summary" on the left taskbar. These reports show the list of schools that were flagged. Summary statistics for each of the schools are also provided. By clicking on the school name, one can navigate to the report for the school.

DATA FILES AND FORMATS

Data files of the AYP I and AYP II results are provided in the folder "Data Files" under the main "PSSA 2009 – Data Forensics" folder. Data files of the student residual and erasure analyses results are also provided. The files are provided by grade level. The file layout of each of the files is provided in the file "2009 Data Forensics File Layout.xls" in the "Data Files" folder.

SUMMARY OF RESULTS

Summary results are provided below. Six tables are provided for each grade. The first table provides the count of flags at the state level. The number of districts[2], schools, and students that

[2] Note that a district is not flagged per se. The flag count for a district indicates that there was a school in that district that was flagged.

were flagged are provided. This is the first table that appears in the HTML State Report. The second table provides the summary statistics of the threat score for each of the indicators. The third table provides the list of schools with three or more flags. This list provides a quick reference to the schools that appear to have the most aberrant testing results. These three tables are for the AYP I analyses. The fourth, fifth, and sixth tables are similar to the prior three tables, except that they provide results for the AYP II analyses.

Grade 3

Table 2. AYP I - Grade 3: Number of Flagged Districts, Schools, and Students.

Flag Counts							
Level	Total	Scale Score	PRO+ADV	Subgroup	Regression	Residual	Erasure
District	58	-	-	-	-	-	-
School	164	34	4	21	23	-	101
Student	652	-	-	-	-	184	474

Table 3. AYP I - Grade 3: Summary Statistics of Threat Score.

	Min	Max	Mean	SD
Math Scale Score	.00	19.60	2.26	2.30
Reading Scale Score	.00	24.90	2.06	2.12
Math Performance Level	.00	13.70	1.73	1.61
Reading Performance Level	.00	9.30	1.68	1.53
Math Regression	.00	14.90	1.26	2.09
Reading Regression	.00	11.90	1.24	2.00
Math ELL	.00	20.20	.84	1.16
Reading ELL	.00	20.20	.83	1.15
Math IEP	.00	9.50	1.46	1.27
Reading IEP	.00	9.00	1.46	1.27
Math ECO	.00	34.90	1.70	2.32
Reading ECO	.00	34.90	1.70	2.31
Math AI/AN	.00	12.80	.31	.61
Reading AI/AN	.00	12.80	.31	.61
Math Asian	.00	7.30	.84	.90
Reading Asian	.00	7.10	.84	.90
Math Black	.00	8.90	1.10	1.05
Reading Black	.00	8.90	1.10	1.06
Math Hispanic	.00	6.00	.96	.94
Reading Hispanic	.00	6.20	.96	.94
Math White	.00	12.80	1.26	1.19
Reading White	.00	12.80	1.26	1.19
Math Multi-Ethnic	.00	6.50	.52	.79
Reading Multi-Ethnic	.00	6.50	.51	.78
Math Erasure	.00	29.10	1.60	3.70
Reading Erasure	.00	35.60	1.25	3.10

Table 4. AYP I - Grade 3: Schools with Three or More Flags.

	School ID	School Name	District ID	District Name	# of Flags
1	000005047	CLIFFORD N PRITTS EL	101261302	CONNELLSVILLE AREA SD	4.00
2	000004819	SOUTH SIDE EL SCH	101261302	CONNELLSVILLE AREA SD	4.00
3	000006930	HEIGHTS TERRACE EL/MIDDLE	118403302	HAZLETON AREA SD	4.00
4	000007340	VALLEY EL/MS	118403302	HAZLETON AREA SD	4.00
5	000007569	WEST HAZLETON EL/MS	118403302	HAZLETON AREA SD	4.00
6	000006534	MCCLURE ALEXANDER K SCH	126515001	PHILADELPHIA CITY SD	4.00
7	000002563	FULTON ELEM SCH	113362603	EPHRATA AREA SD	3.00
8	000007936	HAZLETON ELEM/ MS	118403302	HAZLETON AREA SD	3.00
9	000003763	DISSTON HAMILTON SCH	126515001	PHILADELPHIA CITY SD	3.00
10	000003797	MARSHALL JOHN SCH	126515001	PHILADELPHIA CITY SD	3.00
11	000007522	MUNOZ-MARIN ELEM	126515001	PHILADELPHIA CITY SD	3.00
12	000003750	OLNEY EL SCH	126515001	PHILADELPHIA CITY SD	3.00
13	000002182	BEN FRANKLIN SCH	101268003	UNIONTOWN AREA SD	3.00

Table 5. AYP II - Grade 3: Number of Flagged Districts and Schools.

Flag Counts					
Level	Total	FAY	Scale Score	PRO+ADV	Subgroup
District	31	-	-	-	-
School	94	86	0	0	12

Table 6. AYP II - Grade 3: Summary Statistics of Threat Score.

	Min	Max	Mean	SD
Math Scale Score	.00	1.30	.06	.14
Reading Scale Score	.00	1.40	.06	.14
Math Performance Level	.00	1.50	.07	.16
Reading Performance Level	.00	1.80	.08	.16
Math FAY	.00	99.90	2.59	5.57
Reading FAY	.00	99.90	2.55	5.51

Table 7. AYP II - Grade 3: Schools with Three or More Flags.

	School ID	School Name	District ID	District Name	# of Flags
1	000001013	SNYDER-GIROTTI EL SC	122091303	BRISTOL BOROUGH SD	4.00
2	000007936	HAZLETON ELEM/ MS	118403302	HAZLETON AREA SD	4.00
3	000006930	HEIGHTS TERRACE EL/MIDDLE	118403302	HAZLETON AREA SD	4.00
4	000004463	FORT CRAWFORD SCH	107656303	NEW KENSINGTON-ARNOLD SD	4.00

Grade 4

Table 8. AYP I - Grade 4: Number of Flagged Districts, Schools, and Students.

Flag Counts							
Level	Total	Scale Score	PRO+ADV	Subgroup	Regression	Residual	Erasure
District	63	-	-	-	-	-	-
School	131	47	16	14	30	-	52
Student	749	-	-	-	-	343	412

Table 9. AYP I - Grade 4: Summary Statistics of Threat Score.

	Min	Max	Mean	SD
Math Scale Score	.00	20.90	2.46	2.59
Reading Scale Score	.00	24.50	2.16	2.24
Math Performance Level	.00	17.40	1.88	1.84
Reading Performance Level	.00	18.40	1.87	1.83
Math Regression	.00	25.80	1.34	2.28
Reading Regression	.00	15.90	1.38	2.31
Math ELL	.00	31.70	.84	1.30
Reading ELL	.00	31.70	.83	1.29
Math IEP	.00	9.50	1.45	1.25
Reading IEP	.00	9.50	1.45	1.25
Math ECO	.00	99.90	1.65	2.93
Reading ECO	.00	99.90	1.65	2.93
Math AI/AN	.00	12.50	.31	.59
Reading AI/AN	.00	12.50	.31	.59
Math Asian	.00	9.90	.81	.92
Reading Asian	.00	9.60	.81	.91
Math Black	.00	7.50	1.12	1.07
Reading Black	.00	7.30	1.12	1.06
Math Hispanic	.00	6.10	.95	.96
Reading Hispanic	.00	6.10	.95	.95
Math White	.00	16.70	1.28	1.25
Reading White	.00	16.90	1.28	1.25
Math Multi-Ethnic	.00	6.10	.50	.76
Reading Multi-Ethnic	.00	6.10	.50	.76
Math Erasure	.00	30.50	1.28	3.03
Reading Erasure	.00	36.10	1.02	2.65

Table 10. AYP I - Grade 4: Schools with Three or More Flags.

	School ID	School Name	District ID	District Name	# of Flags
1	000007404	CAYUGA SCHOOL	126515001	PHILADELPHIA CITY SD	5.00
2	000003670	LAMBERTON ROBERT E SCH	126515001	PHILADELPHIA CITY SD	5.00
3	000003797	MARSHALL JOHN SCH	126515001	PHILADELPHIA CITY SD	5.00
4	000002097	BULLSKIN EL SCH	101261302	CONNELLSVILLE AREA SD	4.00
5	000005047	CLIFFORD N PRITTS EL	101261302	CONNELLSVILLE AREA SD	4.00
6	000004819	SOUTH SIDE EL SCH	101261302	CONNELLSVILLE AREA SD	4.00
7	000004444	MONESSEN EL CTR	107655803	MONESSEN CITY SD	4.00
8	000007856	NORTHWOOD ACADEMY CS	126510019	NORTHWOOD ACADEMY CS	4.00
9	000003624	DOUGLASS FREDERICK SCH	126515001	PHILADELPHIA CITY SD	4.00
10	000007521	MARSHALL THURGOOD	126515001	PHILADELPHIA CITY SD	4.00
11	000007233	ZIEGLER WILLIAM H SCH	126515001	PHILADELPHIA CITY SD	4.00
12	000004551	ORE VALLEY EL SCH	112671603	DALLASTOWN AREA SD	3.00
13	000006718	FREELAND EL/MS	118403302	HAZLETON AREA SD	3.00
14	000003602	CATHARINE JOSEPH SCH	126515001	PHILADELPHIA CITY SD	3.00
15	000003708	EDMONDS FRANKLIN S SCH	126515001	PHILADELPHIA CITY SD	3.00
16	000006541	STANTON M HALL SCH	126515001	PHILADELPHIA CITY SD	3.00

Table 11. AYP II - Grade 4: Number of Flagged Districts and Schools.

Flag Counts					
Level	Total	FAY	Scale Score	PRO+ADV	Subgroup
District	36	-	-	-	-
School	86	75	0	0	15

Table 12. AYP II - Grade 4: Summary Statistics of Threat Score.

	Min	Max	Mean	SD
Math Scale Score	.00	2.00	.07	.15
Reading Scale Score	.00	1.90	.06	.14
Math Performance Level	.00	1.80	.07	.16
Reading Performance Level	.00	2.40	.07	.14
Math FAY	.00	99.90	2.50	4.79
Reading FAY	.00	99.90	2.50	5.21

Table 13. AYP II - Grade 4: Schools with Three or More Flags.

	School ID	School Name	District ID	District Name	# of Flags
1	000006930	HEIGHTS TERRACE EL/MIDDLE	118403302	HAZLETON AREA SD	4.00
2	000005350	PAXINOSA EL SCH	120483302	EASTON AREA SD	3.00
3	000001767	MELROSE SCH	115222752	HARRISBURG CITY SD	3.00
4	000002592	HAMILTON EL SCH	113364002	LANCASTER SD	3.00

Grade 5

Table 14. AYP I - Grade 5: Number of Flagged Districts, Schools, and Students.

Flag Counts							
Level	Total	Scale Score	PRO+ADV	Subgroup	Regression	Residual	Erasure
District	102	-	-	-	-	-	-
School	172	48	14	19	64	-	55
Student	646	-	-	-	-	206	440

Table 15. AYP I - Grade 5: Summary Statistics of Threat Score.

	Min	Max	Mean	SD
Math Scale Score	.00	23.60	2.47	2.60
Reading Scale Score	.00	17.30	2.04	2.11
Math Performance Level	.00	14.20	1.98	1.86
Reading Performance Level	.00	14.00	1.87	1.78
Math Regression	.00	31.10	1.57	2.83
Reading Regression	.00	29.10	1.40	2.52
Math ELL	.00	27.70	.81	1.21
Reading ELL	.00	27.60	.80	1.20
Math IEP	.00	11.30	1.44	1.33
Reading IEP	.00	11.20	1.43	1.33
Math ECO	.00	30.70	1.68	2.04
Reading ECO	.00	30.70	1.68	2.04
Math AI/AN	.00	14.30	.30	.63
Reading AI/AN	.00	14.30	.30	.63
Math Asian	.00	5.30	.78	.85
Reading Asian	.00	5.30	.78	.85
Math Black	.00	11.70	1.12	1.07
Reading Black	.00	11.70	1.12	1.07
Math Hispanic	.00	6.80	.90	.92
Reading Hispanic	.00	6.80	.90	.92
Math White	.00	23.00	1.25	1.25
Reading White	.00	23.00	1.25	1.25
Math Multi-Ethnic	.00	5.70	.52	.81
Reading Multi-Ethnic	.00	5.70	.52	.80
Math Erasure	.00	99.90	1.38	3.96
Reading Erasure	.00	99.90	1.12	3.61

Table 16. AYP I - Grade 5: Schools with Three or More Flags.

	School ID	School Name	District ID	District Name	# of Flags
1	000007569	WEST HAZLETON EL/MS	118403302	HAZLETON AREA SD	5.00
2	000005047	CLIFFORD N PRITTS EL	101261302	CONNELLSVILLE AREA SD	4.00
3	000004819	SOUTH SIDE EL SCH	101261302	CONNELLSVILLE AREA SD	4.00
4	000007936	HAZLETON ELEM/ MS	118403302	HAZLETON AREA SD	4.00
5	000007543	ALLIANCE FOR PROGRESS CS	126512990	ALLIANCE FOR PROGRESS CS	3.00
6	000006349	NESCOPECK EL SCH	116191103	BERWICK AREA SD	3.00
7	000007539	CHESTER COMM CS	125232950	CHESTER COMM CS	3.00
8	000006927	GRANDVIEW INTRMD SCH	107651603	DERRY AREA SD	3.00
9	000000169	MOSS SIDE MS	103024102	GATEWAY SD	3.00
10	000007340	VALLEY EL/MS	118403302	HAZLETON AREA SD	3.00
11	000007650	PA CYBER CS	127043430	PA CYBER CS	3.00
12	000007404	CAYUGA SCHOOL	126515001	PHILADELPHIA CITY SD	3.00
13	000007522	MUNOZ-MARIN ELEM	126515001	PHILADELPHIA CITY SD	3.00

Table 17. AYP II - Grade 5: Number of Flagged Districts and Schools.

Flag Counts					
Level	Total	FAY	Scale Score	PRO+ADV	Subgroup
District	35	-	-	-	-
School	85	75	0	0	14

Table 18. AYP II - Grade 5: Summary Statistics of Threat Score.

	Min	Max	Mean	SD
Math Scale Score	.00	1.00	.06	.12
Reading Scale Score	.00	1.50	.06	.13
Math Performance Level	.00	1.10	.07	.13
Reading Performance Level	.00	1.10	.07	.13
Math FAY	.00	99.90	2.59	6.50
Reading FAY	.00	99.90	2.48	5.68

Table 19. AYP II - Grade 5: Schools with Three or More Flags.

	School ID	School Name	District ID	District Name	# of Flags
1	000005292	ALLEN ETHEL DR.	126515001	PHILADELPHIA CITY SD	6.00
2	000001013	SNYDER-GIROTTI EL SC	122091303	BRISTOL BOROUGH SD	4.00
3	000006930	HEIGHTS TERRACE EL/MIDDLE	118403302	HAZLETON AREA SD	4.00
4	000007521	MARSHALL THURGOOD	126515001	PHILADELPHIA CITY SD	4.00

Grade 6

Table 20. AYP I - Grade 6: Number of Flagged Districts, Schools, and Students.

Flag Counts							
Level	Total	Scale Score	PRO+ADV	Subgroup	Regression	Residual	Erasure
District	103	-	-	-	-	-	-
School	166	42	18	23	79	-	45
Student	595	-	-	-	-	105	491

Table 21. AYP I - Grade 6: Summary Statistics of Threat Score.

	Min	Max	Mean	SD
Math Scale Score	.00	36.10	2.64	3.01
Reading Scale Score	.00	24.90	2.15	2.27
Math Performance Level	.00	19.50	2.18	2.32
Reading Performance Level	.00	14.40	1.76	1.76
Math Regression	.00	29.60	1.64	3.27
Reading Regression	.00	28.50	1.63	3.31
Math ELL	.00	18.30	.76	1.15
Reading ELL	.00	18.30	.77	1.16
Math IEP	.00	9.10	1.39	1.34
Reading IEP	.00	9.10	1.39	1.34
Math ECO	.00	26.30	1.81	2.51
Reading ECO	.00	26.30	1.80	2.50
Math AI/AN	.00	14.20	.34	.74
Reading AI/AN	.00	14.20	.34	.73
Math Asian	.00	8.10	.76	.87
Reading Asian	.00	8.40	.76	.87
Math Black	.00	8.70	1.13	1.12
Reading Black	.00	8.70	1.13	1.12
Math Hispanic	.00	6.70	.90	.97
Reading Hispanic	.00	6.70	.90	.96
Math White	.00	21.20	1.22	1.47
Reading White	.00	21.20	1.22	1.46
Math Multi-Ethnic	.00	11.40	.53	.94
Reading Multi-Ethnic	.00	11.20	.53	.94
Math Erasure	.00	99.90	1.49	4.24
Reading Erasure	.00	28.70	1.16	2.77

Table 22. AYP I - Grade 6: Schools with Three or More Flags.

	School ID	School Name	District ID	District Name	# of Flags
1	000007539	CHESTER COMM CS	125232950	CHESTER COMM CS	4.00
2	000005047	CLIFFORD N PRITTS EL	101261302	CONNELLSVILLE AREA SD	4.00
3	000004819	SOUTH SIDE EL SCH	101261302	CONNELLSVILLE AREA SD	4.00
4	000006930	HEIGHTS TERRACE EL/MIDDLE	118403302	HAZLETON AREA SD	4.00
5	000007521	MARSHALL THURGOOD	126515001	PHILADELPHIA CITY SD	4.00
6	000007700	NORTHEAST INTRMD SCH	119357402	SCRANTON SD	4.00
7	000000648	BEAVER FALLS MS	127041503	BIG BEAVER FALLS AREA SD	3.00
8	000006615	CHAMBERSBURG AREA MS	112281302	CHAMBERSBURG AREA SD	3.00
9	000007408	DINGMAN-DELAWARE MS	120522003	DELAWARE VALLEY SD	3.00
10	000000169	MOSS SIDE MS	103024102	GATEWAY SD	3.00
11	000006718	FREELAND EL/MS	118403302	HAZLETON AREA SD	3.00
12	000007569	WEST HAZLETON EL/MS	118403302	HAZLETON AREA SD	3.00
13	000006699	MINERSVILLE AREA EL	129544703	MINERSVILLE AREA SD	3.00
14	000007650	PA CYBER CS	127043430	PA CYBER CS	3.00
15	000007039	STERRETT CLASSICAL ACADEMY	102027451	PITTSBURGH SD	3.00
16	000000856	NORTHWEST MS	114067002	READING SD	3.00
17	000007674	WALTER PALMER LDRSHP LEARNING P	126513490	WALTER PALMER LDRSHP LEARNING PR	3.00
18	000007724	WISSAHICKON CS	126510007	WISSAHICKON CS	3.00

Table 23. AYP II - Grade 6: Number of Flagged Districts and Schools.

Flag Counts					
Level	Total	FAY	Scale Score	PRO+ADV	Subgroup
District	28	-	-	-	-
School	80	74	0	0	10

Table 24. AYP II - Grade 6: Summary Statistics of Threat Score.

	Min	Max	Mean	SD
Math Scale Score	.00	1.00	.07	.14
Reading Scale Score	.00	2.70	.07	.17
Math Performance Level	.00	1.50	.09	.17
Reading Performance Level	.00	1.30	.08	.15
Math FAY	.00	99.90	3.18	8.20
Reading FAY	.00	99.90	3.00	6.84

Table 25. AYP II - Grade 6: Schools with Three or More Flags.

	School ID	School Name	District ID	District Name	# of Flags
1	000006930	HEIGHTS TERRACE EL/MIDDLE	118403302	HAZLETON AREA SD	6.00
2	000007569	WEST HAZLETON EL/MS	118403302	HAZLETON AREA SD	4.00
3	000002595	LINCOLN MS	113364002	LANCASTER SD	4.00
4	000007248	STETSON JOHN B MS	126515001	PHILADELPHIA CITY SD	4.00

Grade 7

Table 26. AYP I - Grade 7: Number of Flagged Districts, Schools, and Students.

Flag Counts							
Level	Total	Scale Score	PRO+ADV	Subgroup	Regression	Residual	Erasure
District	107	-	-	-	-	-	-
School	150	28	21	18	87	-	27
Student	714	-	-	-	-	270	446

Table 27. AYP I - Grade 7: Summary Statistics of Threat Score.

	Min	Max	Mean	SD
Math Scale Score	.00	23.70	2.24	2.60
Reading Scale Score	.00	25.90	2.05	2.38
Math Performance Level	.00	28.80	2.26	2.53
Reading Performance Level	.00	15.30	1.75	1.89
Math Regression	.00	51.30	1.79	4.17
Reading Regression	.00	33.30	1.73	3.72
Math ELL	.00	29.80	.83	1.53
Reading ELL	.00	29.80	.83	1.52
Math IEP	.00	7.80	1.35	1.37
Reading IEP	.00	7.80	1.35	1.37
Math ECO	.00	27.80	1.67	2.38
Reading ECO	.00	27.80	1.66	2.37
Math AI/AN	.00	3.40	.34	.54
Reading AI/AN	.00	3.40	.34	.54
Math Asian	.00	6.60	.78	.94
Reading Asian	.00	6.60	.78	.94
Math Black	.00	8.10	1.12	1.17
Reading Black	.00	8.00	1.12	1.16
Math Hispanic	.00	7.80	.92	1.01
Reading Hispanic	.00	7.70	.92	1.01
Math White	.00	11.00	1.18	1.27
Reading White	.00	11.30	1.18	1.27
Math Multi-Ethnic	.00	8.00	.55	.94
Reading Multi-Ethnic	.00	8.00	.54	.94
Math Erasure	.00	99.90	1.42	4.40
Reading Erasure	.00	39.90	1.20	3.13

Table 28. AYP I - Grade 7: Schools with Three or More Flags.

	School ID	School Name	District ID	District Name	# of Flags
1	000003835	WAGNER GEN LOUIS MS	126515001	PHILADELPHIA CITY SD	6.00
2	000003824	THEODORE ROOSEVELT MS	126515001	PHILADELPHIA CITY SD	5.00
3	000007539	CHESTER COMM CS	125232950	CHESTER COMM CS	3.00
4	000007340	VALLEY EL/MS	118403302	HAZLETON AREA SD	3.00
5	000007650	PA CYBER CS	127043430	PA CYBER CS	3.00
6	000006913	FITLER ACADEMICS PLUS	126515001	PHILADELPHIA CITY SD	3.00
7	000003781	FITZPATRICK ALOYSIUS	126515001	PHILADELPHIA CITY SD	3.00
8	000006641	MARKLE INTRMD SCH	112676403	SOUTH WESTERN SD	3.00

Table 29. AYP II - Grade 7: Number of Flagged Districts and Schools.

Flag Counts					
Level	Total	FAY	Scale Score	PRO+ADV	Subgroup
District	44	-	-	-	-
School	92	82	0	0	15

Table 30. AYP II - Grade 7: Summary Statistics of Threat Score.

	Min	Max	Mean	SD
Math Scale Score	.00	2.70	.09	.20
Reading Scale Score	.00	3.80	.08	.20
Math Performance Level	.00	3.30	.11	.22
Reading Performance Level	.00	2.10	.10	.20
Math FAY	.00	99.90	3.68	9.46
Reading FAY	.00	99.90	3.69	9.81

Table 31. AYP II - Grade 7: Schools with Three or More Flags.

	School ID	School Name	District ID	District Name	# of Flags
1	000007936	HAZLETON ELEM/ MS	118403302	HAZLETON AREA SD	4.00
2	000006930	HEIGHTS TERRACE EL/MIDDLE	118403302	HAZLETON AREA SD	4.00
3	000002600	WHEATLAND MS	113364002	LANCASTER SD	4.00
4	000007676	CEP - HUNTING PARK	126515001	PHILADELPHIA CITY SD	4.00
5	000007663	GROVER WASHINGTON MS	126515001	PHILADELPHIA CITY SD	3.00

Grade 8

Table 32. AYP I - Grade 8: Number of Flagged Districts, Schools, and Students.

Flag Counts							
Level	Total	Scale Score	PRO+ADV	Subgroup	Regression	Residual	Erasure
District	108	-	-	-	-	-	-
School	155	36	13	17	95	-	27
Student	593	-	-	-	-	136	459

Table 33. AYP I - Grade 8: Summary Statistics of Threat Score.

	Min	Max	Mean	SD
Math Scale Score	.00	36.10	2.42	3.01
Reading Scale Score	.00	28.50	2.06	2.42
Math Performance Level	.00	29.20	2.01	2.28
Reading Performance Level	.00	27.20	1.71	1.89
Math Regression	.00	36.10	1.77	3.81
Reading Regression	.00	37.50	1.85	4.17
Math ELL	.00	20.70	.79	1.21
Reading ELL	.00	20.70	.79	1.21
Math IEP	.00	13.00	1.47	1.52
Reading IEP	.00	13.00	1.47	1.51
Math ECO	.00	25.20	1.71	2.30
Reading ECO	.00	25.00	1.72	2.30
Math AI/AN	.00	3.50	.38	.58
Reading AI/AN	.00	3.50	.37	.57
Math Asian	.00	7.00	.78	.93
Reading Asian	.00	7.00	.77	.93
Math Black	.00	12.30	1.11	1.17
Reading Black	.00	11.90	1.11	1.17
Math Hispanic	.00	6.20	.95	1.04
Reading Hispanic	.00	6.20	.95	1.05
Math White	.00	9.50	1.17	1.20
Reading White	.00	9.50	1.17	1.20
Math Multi-Ethnic	.00	8.30	.50	.89
Reading Multi-Ethnic	.00	8.30	.50	.89
Math Erasure	.00	99.90	1.38	4.58
Reading Erasure	.00	99.90	1.26	4.41

Table 34. AYP I - Grade 8: Schools with Three or More Flags.

	School ID	School Name	District ID	District Name	# of Flags
1	000006925	SHAWNEE MS	120483302	EASTON AREA SD	4.00
2	000003824	THEODORE ROOSEVELT MS	126515001	PHILADELPHIA CITY SD	4.00
3	000004329	BELLMAR MS	107650603	BELLE VERNON AREA SD	3.00
4	000007539	CHESTER COMM CS	125232950	CHESTER COMM CS	3.00
5	000007569	WEST HAZLETON EL/MS	118403302	HAZLETON AREA SD	3.00
6	000007777	MARITIME ACADEMY CS	126510014	MARITIME ACADEMY CS	3.00
7	000003883	NORTH SCHUYLKILL JSHS	129545003	NORTH SCHUYLKILL SD	3.00
8	000007650	PA CYBER CS	127043430	PA CYBER CS	3.00
9	000007244	PENN TREATY MS	126515001	PHILADELPHIA CITY SD	3.00
10	000003835	WAGNER GEN LOUIS MS	126515001	PHILADELPHIA CITY SD	3.00
11	000007253	WILSON WOODROW MS	126515001	PHILADELPHIA CITY SD	3.00

Table 35. AYP II - Grade 8: Number of Flagged Districts and Schools.

Flag Counts					
Level	Total	FAY	Scale Score	PRO+ADV	Subgroup
District	48	-	-	-	-
School	88	79	0	0	10

Table 36. AYP II - Grade 8: Summary Statistics of Threat Score.

	Min	Max	Mean	SD
Math Scale Score	.00	1.80	.09	.15
Reading Scale Score	.00	1.50	.08	.15
Math Performance Level	.00	1.60	.11	.19
Reading Performance Level	.00	2.10	.11	.22
Math FAY	.00	99.90	3.60	8.49
Reading FAY	.00	99.90	3.56	8.44

Table 37. AYP II - Grade 8: Schools with Three or More Flags.

	School ID	School Name	District ID	District Name	# of Flags
1	000002754	LEBANON MS	113384603	LEBANON SD	4.00
2	000008029	BARRY COMM JOHN SCH	126515001	PHILADELPHIA CITY SD	4.00

Grade 11

Table 38. AYP I - Grade 11: Number of Flagged Districts, Schools, and Students.

Flag Counts							
Level	Total	Scale Score	PRO+ADV	Subgroup	Regression	Residual	Erasure
District	105	-	-	-	-	-	-
School	135	40	25	24	68	-	25
Student	453	-	-	-	-	123	330

Table 39. AYP I - Grade 11: Summary Statistics of Threat Score.

	Min	Max	Mean	SD
Math Scale Score	.00	29.40	2.37	3.03
Reading Scale Score	.00	32.60	2.40	3.17
Math Performance Level	.00	21.30	2.13	2.52
Reading Performance Level	.00	23.80	1.91	2.30
Math Regression	.00	61.70	1.60	4.24
Reading Regression	.00	62.50	1.64	3.84
Math ELL	.00	10.30	.74	1.17
Reading ELL	.00	12.40	.74	1.20
Math IEP	.00	20.40	1.42	1.64
Reading IEP	.00	20.10	1.42	1.63
Math ECO	.00	99.90	2.32	6.43
Reading ECO	.00	99.90	2.31	6.43
Math AI/AN	.00	3.30	.39	.62
Reading AI/AN	.00	3.30	.39	.62
Math Asian	.00	5.90	.81	.94
Reading Asian	.00	5.80	.81	.94
Math Black	.00	28.20	1.12	1.46
Reading Black	.00	28.10	1.11	1.45
Math Hispanic	.00	8.10	.90	1.05
Reading Hispanic	.00	8.50	.89	1.05
Math White	.00	9.10	1.21	1.30
Reading White	.00	9.80	1.20	1.29
Math Multi-Ethnic	.00	7.00	.43	.79
Reading Multi-Ethnic	.00	7.00	.42	.79
Math Erasure	.00	99.90	1.43	5.92
Reading Erasure	.00	35.50	1.23	3.24

Table 40. AYP I - Grade 11: Schools with Three or More Flags.

	School ID	School Name	District ID	District Name	# of Flags
1	000007542	IMHOTEP INSTITUTE CS	126512980	IMHOTEP INSTITUTE CS	7.00
2	000007578	ARCH AND DESIGN CHS	126513190	ARCH AND DESIGN CHS	6.00
3	000001892	PENN WOOD SHS	125239652	WILLIAM PENN SD	6.00
4	000003850	FRANKFORD HS	126515001	PHILADELPHIA CITY SD	5.00
5	000003855	NORTHEAST HS	126515001	PHILADELPHIA CITY SD	5.00
6	000006413	SCRANTON HS	119357402	SCRANTON SD	5.00
7	000003260	CHELTENHAM HS	123461302	CHELTENHAM TOWNSHIP SD	4.00
8	000006643	CONNELLSVILLE AREA CTC	101266007	CONNELLSVILLE AREA CTC	4.00
9	000007737	PHILA ELEC & TECH CHS	126510009	PHILA ELEC & TECH CHS	4.00
10	000007255	STRAWBERRY MANSION HS	126515001	PHILADELPHIA CITY SD	4.00
11	000003210	PLEASANT VALLEY HS	120455203	PLEASANT VALLEY SD	4.00
12	000005287	STRATH HAVEN HIGH SCHOOL	125239603	WALLINGFORD-SWARTHMORE SD	4.00
13	000005091	ABINGTON HEIGHTS HS	119350303	ABINGTON HEIGHTS SD	3.00
14	000006555	AMBRIDGE AREA HS	127040703	AMBRIDGE AREA SD	3.00
15	000004957	FREEDOM HS	120481002	BETHLEHEM AREA SD	3.00
16	000000802	GOVERNOR MIFFLIN SHS	114063003	GOVERNOR MIFFLIN SD	3.00
17	000007348	HAZLETON AREA HS	118403302	HAZLETON AREA SD	3.00
18	000007650	PA CYBER CS	127043430	PA CYBER CS	3.00
19	000004682	PENNSBURY SHS	122098202	PENNSBURY SD	3.00
20	000000859	READING SHS	114067002	READING SD	3.00
21	000003367	SPRING-FORD SHS	123467303	SPRING-FORD AREA SD	3.00
22	000004080	WELLSBORO AREA HS	117598503	WELLSBORO AREA SD	3.00

Table 41. AYP II - Grade 11: Number of Flagged Districts and Schools.

Flag Counts					
Level	Total	FAY	Scale Score	PRO+ADV	Subgroup
District	53	-	-	-	-
School	69	67	0	0	5

Table 42. AYP II - Grade 11: Summary Statistics of Threat Score.

	Min	Max	Mean	SD
Math Scale Score	.00	2.00	.13	.21
Reading Scale Score	.00	2.60	.12	.21
Math Performance Level	.00	2.30	.10	.18
Reading Performance Level	.00	2.10	.11	.19
Math FAY	.00	99.90	3.60	10.83
Reading FAY	.00	99.90	3.48	10.30

Table 43. AYP II - Grade 11: Schools with Three or More Flags.

	School ID	School Name	District ID	District Name	# of Flags
1	000003816	FELS SAMUEL HS	126515001	PHILADELPHIA CITY SD	4.00
2	000007899	ACTS @ WILLIAM PENN SCHOOL	115222752	HARRISBURG CITY SD	3.00
3	112000000	LINCOLN IU 12	112000000	LINCOLN IU 12	3.00

Chapter Four: Discussion

The purpose of the data forensic analyses was to identify schools and students who may have engaged in testing behavior that may have allowed them to gain an unfair advantage. With test scores that may not have been earned fairly, valid inferences from those scores cannot be made. Students were identified if their response pattern did not match that predicted by the Rasch model (i.e., the residual analysis) or if they had a large number of wrong-to-right erasures compared to the state mean. Schools were identified if they had a large change in scale score, percentage of proficient or advanced students, difference between their actual and predicted mean scale score, number of wrong-to-right erasures compared to the state mean, and subgroup participation rates.

Again, it should be noted that the schools and students were flagged based on statistical evidence alone. Their scores, response pattern, and number of erasures were aberrant, from a statistical probability perspective. This does not imply that the school or student engaged in inappropriate testing activity. The statistical evidence merely suggests that something aberrant or unusual occurred.

In this first year of data forensic analyses for the PSSA, these results should be used with caution. Lists of schools that were flagged by multiple indicators were provided above. This list serves as a good starting point when examining schools and students for potential testing irregularities. Of particular interest are the schools that were flagged by the erasure analysis indicator. The schools identified by this indicator had a very large number of wrong-to-right erasures compared to the state mean. Note that their total number of erasures was not much higher than the number of wrong-to-right erasures. Thus, most of their erasures were of the wrong-to-right type. These results may strongly suggest that a testing irregularity occurred at the school. Reports in HTML format are provided for the schools that were flagged by the erasure analysis as well as the other indicators. The details in these reports may help PDE resolve any inquiries about a school's test results during the AYP appeals process.

References

Agresti, A. (2002). *Categorical Data Analysis (2nd Edition)*. New York: Wiley.

Gart, J. J., and Zweifer, J. R. (1967). On the bias of various estimators of the logit and its variance with applications to quantal bioassay. *Biometrika*, 54, 181-187.

Mead, R. (1976). *Analysis of fit of data to the Rasch model with residuals*. Unpublished dissertation. University of Chicago.

Mead, R. (1980). Using the Rasch model to identify person-based measurement disturbances. *Proceedings of the 1979 Computer Adaptive Testing Conference*. Minneapolis, MN: University of Minnesota.

No Child Left Behind Act of 2001, Pub. L. No. 107–110, 115 Stat. 1425 (2002).

Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen: Danish Institute for Educational Research.

Smith, R. (1982). Detecting measurement disturbance with the Rasch model. Unpublished dissertation. University of Chicago.

Smith, R. (1986). Person fit in Rasch model. *Educational and Psychological Measurement*, 46, 359-372.

Smith, R. (2004). Fit analysis in latent trait measurement models in Smith & Smith, (Eds.) *Introduction to Rasch measurement.* Chapter 4, 73-92.