

**TESTING INTEGRITY SYMPOSIUM**  
**Issues and Recommendations for Best Practice**

**U.S. DEPARTMENT OF EDUCATION**  
**Institute of Education Sciences**  
**National Center for Education Statistics**  
**2013**

**Testing Integrity Symposium  
Issues and Recommendations for Best Practice**

**Table of Contents**

|   |            |
|---|------------|
| <b>SYMPOSIUM PANELISTS</b>  | <b>iii</b> |
| <b>INTRODUCTION</b>   | <b>1</b>   |
| <b>SECTION I. PREVENTION OF IRREGULARITIES IN ACADEMIC TESTING</b>  | <b>2</b>   |
| <b>SECTION II. DETECTION AND ANALYSIS OF IRREGULARITIES IN ACADEMIC TESTING</b>                               | <b>8</b>   |
| <b>SECTION III. RESPONSE AND INVESTIGATION OF ALLEGED AND/OR ACTUAL TESTING<br/>IRREGULARITIES</b>            | <b>10</b>  |
| <b>SECTION IV. TESTING INTEGRITY PRACTICES AND PROCEDURES FOR ONLINE AND<br/>TECHNOLOGY-BASED ASSESSMENTS</b> | <b>15</b>  |
| <b>CLOSING COMMENTS</b>   | <b>18</b>  |
| <b>REFERENCES CITED</b>   | <b>20</b>  |
| <b>APPENDIX A: TESTING INTEGRITY SYMPOSIUM PANELIST BIOGRAPHIES</b>   | <b>21</b>  |
| <b>APPENDIX B: REQUEST FOR INFORMATION (RFI) RESPONSES</b>  | <b>25</b>  |

## Symposium Panelists

### **Tony Alpert**

Chief Operating Officer  
SMARTER Balanced Assessment Consortium

### **Audrey Amrein-Beardsley**

Associate Professor  
Arizona State University

### **Wes Bruce**

Chief Assessment Officer  
Indiana Department of Education/PARCC

### **Wayne Camara**

Vice President for Research and Development  
College Board

### **Gregory J. Cizek**

Professor of Educational Measurement and  
Evaluation  
University of North Carolina at Chapel Hill

### **Tisha S. Edwards**

Chief of Staff  
Baltimore City Public Schools

### **Lou Fabrizio**

Director of Data, Research and Federal Policy  
North Carolina Department of Public Instruction

### **Steve Ferrara**

Vice President  
Performance Assessment, Pearson

### **Melissa Fincher**

Associate Superintendent, Assessment and  
Accountability  
Georgia Department of Education

### **David Foster**

Chairman and CEO, Caveon Consulting Services  
Caveon, LLC

### **John Fremer**

President, Caveon Consulting Services  
Caveon, LLC

### **Brian Jacob**

Walter H. Annenberg Professor of Education Policy  
University of Michigan

### **James Liebman**

Simon H. Rifkind Professor of Law  
Columbia University Law School

### **Scott Norton**

Strategic Initiative Director, Standards, Assessment,  
and Accountability  
Council of Chief State School Officers

### **Carswell Whitehead**

Test Security Director Office of Testing Integrity,  
Educational Testing Service

### **Bob Wilson**

Partner  
Wilson, Morton & Downs, LLC

Note: The views expressed in this document are those of parties outside the U.S. Department of Education and do not necessarily represent the views of the Department. The information, which is provided as a service to the general public, school officials, and educators, describes practices that may contribute to the integrity of testing procedures and results. School officials, educators, and the public may find these suggestions helpful in improving test administration practices. The Department does not require or endorse the use of these practices, and recognizes that no single particular approach is appropriate in all circumstances.

*Recent news reports of widespread or suspected cheating on standardized tests in several school districts around the country have been taken by some as evidence that we must reduce reliance on testing to measure student growth and achievement. Others have gone even farther, claiming that cheating is an inevitable consequence of “high-stakes testing” and that we should abandon testing altogether. To be sure, there are lessons to be learned from these jarring incidents, but the existence of cheating says nothing about the merits of testing. Instead, cheating reflects a willingness to lie at children’s expense to avoid accountability—an approach I reject entirely.*

**– U.S. Secretary of Education Arne Duncan, *The Washington Post*, July 19, 2011**

## Introduction

Educators, parents, and the public depend on accurate, valid, reliable, and timely information about student academic performance. The availability of test data is important to improve instruction, identify the needs of individual students, implement targeted interventions, and help all students reach high levels of achievement. Testing irregularities – breaches of test security or improper administration of academic testing – undermine efforts to use those data to improve student achievement. Unfortunately, there have been high-profile and systemic incidents of cheating in several school districts across the country in recent years.

While every state has policies in place to address test administration, no “library of best practices” exists that could help state educational agencies (SEAs) and local educational agencies (LEAs) prevent, detect, and respond to irregularities in academic testing. In light of the recent reports of misconduct by school officials in the test administration process, and the importance of that process, the U. S. Department of Education (Department) sought to collect and share information about practices and policies that have been used to prevent, detect, and respond to irregularities in academic testing.

The Department published a request for information (RFI) in the *Federal Register*<sup>1</sup> on January 17, 2012, asking the public to submit best practices and policies regarding the prevention, detection, and investigation of irregularities in academic testing. The Department received 19 responses from a variety of sources, including academic researchers, testing companies, SEAs, law firms, and nonprofit organizations.

In addition to the RFI, the Institute of Education Sciences’ National Center for Education Statistics (NCES) sponsored a Testing Integrity Symposium (Symposium) in Washington, D.C. on February 28, 2012.<sup>2</sup> At the day-long Symposium 16 experts from across the nation participated in a series of four panels to share and discuss best practices regarding the prevention, detection, and investigation of irregularities in academic testing, and how these best practices might change for assessments delivered online and by computer. The panelists included state and local school officials, academic researchers, and members of the professional testing community.<sup>3</sup> Chancellor Kaya Henderson of the Washington, D.C. Department

<sup>1</sup> Federal Register Vol. 77, No. 10 / Tuesday, January 17, 2012 / Notices: <http://www.gpo.gov/fdsys/pkg/FR-2012-01-17/pdf/2012-753.pdf>.

<sup>2</sup> General information about the Symposium can be found at the following Internet address: <http://ies.ed.gov/whatsnew/conferences/?id=966&cid=2>.

<sup>3</sup> Information about panelists is included in Appendix A, Testing Integrity Symposium Panelist Biographies.

of Public Schools and Kathi M. King, a teacher at Messalonskee High School in Oakland, Maine made opening remarks about the value of testing and the impact of testing irregularities on school administrators and classroom teachers. Jack Buckley, the commissioner of NCES, moderated the four panels. The Symposium was open to the public and broadcast online. Nearly 90 participants attended in person and more than 400 participants viewed the proceedings live on the Internet via webcast.

### **Sources and purpose of this report**

This report draws upon three sources of information about practices that support the integrity of test results: the opinions of experts and practitioners as expressed in the RFI responses, the comments and discussions from the Symposium, and, where available, policy manuals or professional standards published by SEAs and professional associations. It is organized by topic, summarizing practices and policies related to four areas of testing integrity presented at the Symposium. Each section includes information from all three sources listed above. Symposium participants' comments are presented where they are most relevant to a topic and may not follow the exact order of presentation.<sup>4</sup> RFI responses are similarly included under applicable content areas. Appendix B, Request for Information (RFI) Responses, lists the individuals and agencies who replied to this request.

The RFI and Symposium are part of a broader effort by the Department to identify and disseminate practices and policies to SEAs, LEAs, and the testing companies that can assist them in their continuing efforts to improve the validity and reliability of assessment results. This report consists largely of the opinions of experts who presented at the Symposium or responded to the RFI. The Department hopes that this document will be a starting point for further dialogue around the integrity of academic assessments and that it will help SEAs and LEAs identify, share, and implement best practices for preventing, detecting, and investigating irregularities in testing. The practices and policies summarized in this report reflect the expertise and opinions of outside experts and education practitioners. They do not represent endorsements by the Department or the Department's official position on these matters.

As was the case with the RFI and the Symposium, this summary focuses on four areas related to testing integrity: (1) the prevention of irregularities in academic testing; (2) the detection and analysis of testing irregularities; (3) the response to an investigation of alleged and/or actual misconduct; and (4) testing integrity practices for technology-based assessments.

## **Section I. Prevention of Irregularities in Academic Testing**

This section provides experts' insight on the issue of preventing testing irregularities. It focuses on best practices and policies that SEAs and LEAs have implemented to prevent testing irregularities; barriers to implementing those practices and policies; and the role school culture plays in testing security.

### **Develop a definition of cheating.**

According to panelist Dr. Amrein-Beardsley, an important first step is establishing a common definition of cheating in the context of academic testing. Dr. Amrein-Beardsley stated that there are varying degrees of cheating, making it difficult to quantify its incidence. She developed

---

<sup>4</sup> See <[http://nces.ed.gov/whatsnew/conferences/PDF/2012\\_symposium\\_transcript.pdf](http://nces.ed.gov/whatsnew/conferences/PDF/2012_symposium_transcript.pdf)> for a complete transcript of the symposium. Subsequent footnotes will identify this source as "transcript."

taxonomy in order to understand what practices can lead to testing irregularities and to prevent testing irregularities from occurring in the future.<sup>5</sup> Based on survey research, this taxonomy of cheating, which includes first-, second-, and third-degree offenses depending upon severity, is modeled on the legal classification of crimes. Although the Symposium panelists discussed all three different degrees of cheating, ranging from involuntary and accidental to willful and premeditated, this summary focuses on the intentional practices included in the taxonomy's first and second degrees.<sup>6</sup>

Cheating in the first degree refers to willful and sometimes premeditated acts including:

- Erasing and changing students' answers;
- Filling in answers left blank by students;
- Overtly and covertly providing correct answers on tests;
- Falsifying student test identification or tracking numbers; and
- Suspending or otherwise excluding students with poor academic performance on testing days, so that they are not tested.

Cheating in the second degree includes more subtle forms of misconduct such as:

- Cueing students on incorrect answers (for example, tapping on the desk or nudging);
- Distributing "cheat-sheets," talking students through processes and definitions; and
- Giving extra time on tests during recess or before/after school.<sup>7,8</sup>

### **Establish a healthy testing culture.**

The panelists emphasized the importance of preventing cheating from occurring in the first place by creating a culture in which tests are focused on the students and integrity is ingrained in the school district's culture. According to Mr. Wilson, first-, second-, and third-degree cheating is the foreseeable result of a testing culture that relies upon on pressure and intimidation to meet unreasonable academic targets.<sup>9</sup> Mr. Liebman noted that it was important to "develop a culture in the classroom that the data is really useful and helpful ... then the

---

<sup>5</sup> Transcript p. 33.

<sup>6</sup> See Amrein-Beardsley, Audrey, Berliner, David C., Rideau, Sharon, *Cheating in the first, second, and third degree: Educators' responses to high-stakes testing*, Education Policy Analysis Archives (2010).

Volume: 18, Issue: 14, Pages: 1-36.

<sup>7</sup> Transcript pp. 33-5.

<sup>8</sup> This paper does not address practices for preventing, detecting, and addressing "cheating in the third degree," which Dr. Amrein-Beardsley defined as generally involuntary and unintentional. In many of these situations, teachers do not believe they are "cheating," but may in fact believe their actions are good test preparation for their students. Examples of cheating in the third degree include (1) "teaching to the test" by using previous academic tests or having access to blueprints; (2) "narrowing the curriculum" by excluding or postponing other educationally relevant subjects leading up to academic testing; and (3) focusing inordinately on test taking strategies.

<sup>9</sup> Transcript pp. 208-9.

reliability of that data becomes important, including to everybody who's using it, and then ... educators understand [that] reliability can become a problem for them."<sup>10</sup>

The panelists agreed that the superintendent of a school system and other leaders should "set the tone" for acting responsibly regarding testing practices. School leaders should encourage and embrace an honor code for all educators that outlines the importance of integrity during test administration and throughout the year. Educators should understand the importance of academic testing and how irregularities can damage the school, community, and students, and should be "responsible for implementing a culture where learning is the goal, as opposed to performance."<sup>11</sup>

### **Focus on high-risk threats first.**

Since LEAs and SEAs have limited resources and time, focusing on high-risk, high-probability threats is the most cost-effective approach to averting the most damaging threats to testing integrity. According to Mr. Foster, a high-risk, high-probability threat is defined as immediate, likely to occur, and very damaging.<sup>12</sup> High-risk threats include the sharing of answers, exposure of actual test items prior to administration, or proxy test-taking.<sup>13</sup>

### **Train and certify principals and teachers in administering and interpreting academic assessments.**

Professor Cizek stressed that proper training and professional development at all levels is crucial in creating a healthy testing culture. Principals and teachers should be properly trained, prepared, and qualified in administering and interpreting academic assessments.<sup>14</sup> According to Mr. Norton, because of increased incentives for adults to cheat on assessments "what has been adequate in the past [regarding teacher training] is soon going to be inadequate."<sup>15</sup> Teachers should be trained to see testing as an important professional responsibility, and testing should be viewed as an opportunity to let "students shine."<sup>16</sup>

Nonetheless, according to Dr. Cizek, many states do not require principals and teachers to be qualified in administering assessments. This lack of training sometimes "leads to a distrust of testing," and a misunderstanding about the reliability and validity of standardized testing in general.<sup>17</sup> This lack of trust in assessments exists even though state academic assessments are generally "far and away the most objective, fairest, least biased, most dependable measure that [students] will see the entire year," according to Cizek.<sup>18</sup>

In some states, such as Oregon, LEAs "train school test coordinators annually on test administration and test security expectations and best practices." This training covers the test administration and security requirements described in Oregon's test administration manual, and "serves to remind test coordinators of test security protocol and the procedures for preventing

---

<sup>10</sup> Transcript p. 77.

<sup>11</sup> Transcript pp. 33-5, 40, 42.

<sup>12</sup> Transcript p. 114.

<sup>13</sup> John Fremer for Caveon, Response to US DOE RFI to Gather Technical Expertise Pertaining to Testing Integrity, p.

10.

<sup>14</sup> Transcript pp. 44-5.

<sup>15</sup> Transcript p. 70.

<sup>16</sup> Transcript p. 67.

<sup>17</sup> Transcript p.43.

<sup>18</sup> Transcript p. 44.

and responding to test improprieties.” After the initial training, LEA test coordinators provide ongoing support to test administrators throughout the LEA “by distributing testing-related updates and reminders, answering questions about appropriate testing practices, providing refresher trainings as needed, and periodically conducting drop-in visits to observe testing sites in their district.”<sup>19</sup> Mr. Liebman stated that in order to address instances of first- and second-degree cheating New York City assigns a test coordinator to each school. The coordinator has primary responsibility for test security, works with the principal in test administration, and prevents too much pressure from being placed on the principal.

These practices and recommendations agree with guidelines prepared by the National Council on Measurement in Education (NCME) stating that “training should provide an overview of ethical and proper administration procedures and stress the importance of academic and assessment integrity as a means of avoiding serious negative consequences for the testing program and its potential damage to the educational reputation of students and schools. Staff and students should understand and support monitoring efforts to report and detect breaches of security, cheating, and other improper behavior.”<sup>20</sup>

#### **Develop standard policies and procedures for test administration.**

Panelists emphasized that clear policies, procedures, and protocols regarding test administration are essential to prevent misconduct. Strong and comprehensive language should communicate clearly to staff instructions for test administration, procedures for secure management of testing materials, protocols for reporting breaches (*e.g.*, anonymous tip hotlines and other reporting systems), explanations of methods used to detect irregularities, and sanctions for misconduct. Misconduct is less likely to occur if staff members know that a comprehensive policy and system is in place.<sup>21</sup> For example, in the New York City school system any personnel who are involved in the test administration process must sign a statement confirming that they have read the testing handbook and understand the penalties for violating the rules therein.<sup>22</sup>

The Oregon Department of Education (ODE) has developed uniform test administration requirements that all Oregon LEAs “must follow when administering an Oregon statewide assessment to ensure both test reliability and validity from classroom to classroom, teacher to teacher, school to school, and district to district.” Oregon’s test administration requirements address topics such as the “test environment, the interactions between the test administrator and students, the resources students may access during testing, and the accommodations students may receive during testing.” The ODE publishes these requirements in a test administration manual and also codifies them in the state’s administrative code. State law requires that all school and district staff involved in the administration of statewide assessments know and adhere to the policies and procedures included in the manual.

---

<sup>19</sup> Oregon Department of Education Testing Integrity [RFI] Response, p. 2, February 13, 2012.

<sup>20</sup> National Council on Measurement in Education (2012). Testing and Data Integrity in the Administration of Statewide Student Assessment Programs, p. 4, October 2012. Available online at [http://ncme.org/default/assets/File/Committee%20Docs/Test%20Score%20Integrity/Test%20Integrity-NCME%20Endorsed%20\(2012%20FINAL\).pdf](http://ncme.org/default/assets/File/Committee%20Docs/Test%20Score%20Integrity/Test%20Integrity-NCME%20Endorsed%20(2012%20FINAL).pdf). This report uses the October 2012 publication rather than the unpublished working draft included in NCME’s response to the January 2012 RFI.

<sup>21</sup> Transcript p. 61.

<sup>22</sup> Transcript p. 61.



According to the ODE, by “establishing clear expectations for statewide test administration practices” it has limited the amount of interpretation individual test administrators must make when administering statewide assessments. This standardization of testing practices has “led to more consistent practices between classrooms and across districts” and has also led to a “heightened awareness at the district level of what practices are considered test improprieties.”<sup>23</sup>

The NCME guidelines regarding testing integrity state that school personnel should have an opportunity to provide input on the development of these policies and procedures, and “be given ample lead time for implementation before the policy becomes effective.”<sup>24</sup>

**Keep testing windows short.**

Panelists agreed that short testing windows (*i.e.*, all students should be taking the test at the same time or close to the same time as possible) reduce the probability that testing booklets and other types of information such as test items will be shared inappropriately with students taking the test at a later time.

**Administer tests in controlled environments.**

Tests should be administered in controlled and secure environments that limit access to curricular materials, resources, and other visuals that could aid students. Any resources that students could use as cues or triggers should not be present during test administration. Although such a practice would appear to be intuitive, Dr. Amrein-Beardsley cited a survey in which many teachers responded that “they thought it was not fair to make the environment artificial,” and that “students should have access to those resources” because they are available in the students’ day-to-day environment.<sup>25</sup>

**Establish and monitor the chain of custody.**

Several panelists emphasized the importance of establishing a chain of custody, and retaining tight control over testing materials to prevent tampering. In general, the panelists emphasized that test administrators should:

- Limit who has access to the materials;
- Make sure all who have contact with the testing materials record when and where they accessed the materials;
- Ensure the materials are kept in a secure location when not being administered (and document who has access to the secure location);
- Make clear that failure to secure materials or providing false or misleading information on the chain of custody may carry consequences; and

---

<sup>23</sup> Oregon Department of Education Testing Integrity [RFI] Response, pp. 1-2, February 13, 2012. Oregon’s Test Administration Manual is available online at <http://www.ode.state.or.us/go/tam>.

<sup>24</sup> National Council on Measurement in Education (2012). Testing and Data Integrity in the Administration of Statewide Student Assessment Programs, p. 4.

<sup>25</sup> Transcript pp. 38-9.

- Note how long chain of custody documentation needs to be maintained by the school.

In Oregon LEAs help prevent test improprieties by controlling which staff members have access to the statewide testing system and confidential student test data. One promising practice in Oregon has been for “school test coordinators to track which staff have received annual test administration and security training and signed an Assurance of Test Security form, and to only set up user accounts in the testing system for those staff who have met these requirements.”<sup>26</sup>

Testing booklets should be pre-packaged (e.g., shrink wrapped) before they arrive at testing sites, and should be marked with pre-determined student identification numbers. Test coordinators should have a process for establishing a chain of custody as described above for the testing booklets so that they can be tracked.<sup>27</sup>

Because it is vital for security purposes to know who administered a test, the “simple addition to a test developer’s and state’s procedures of including batch headers is a really important step.”<sup>28</sup> Batch headers are serial-numbered documents that indicate who monitored the answer sheets contained in the batch with the corresponding serial number. Test administrators can use batch headers to identify who monitored the answer sheets that are processed by scanning and scoring systems.

**Remove testing materials from the testing location immediately and score them off-site.**

Several panelists recommended that school officials should remove testing materials from the testing location immediately following test administration and score tests off-site to prevent tampering with answer sheets. For testing items that cannot be scored by computer, but must be graded by teachers, tests should be taken to a different school. Tests should be scored by at least two or three teachers with an assigned “table leader” comparing the scores to determine whether the test results from different scorers agree. This will ensure inter-rater reliability and remove the opportunity for any one examiner to tamper with answer sheets.<sup>29</sup>

New York City is looking for innovative ways improve its post-test administration that would further prevent cheating. Specifically, the school system is determining whether it is feasible to have “somebody come in and do all of the distribution, picking up, everything scanned, including the constructed responses, so they go immediately into the machines and can’t be tampered with.” These tests could be distributed “wherever you want them to be graded,” and would allow for instant, automatic, city-wide comparison of how each score compares to scores elsewhere in the city” to ensure inter-rater reliability.<sup>30</sup>

**SEAs should monitor test administration.**

RFI responders and Symposium panelists agreed that the SEA should conduct some level of monitoring of the test administration process and should require that its LEAs monitor test administration in their schools. In Maryland, for example, the Maryland State Department of

---

<sup>26</sup> Oregon Department of Education Testing Integrity [RFI] Response, p. 5, February 13, 2012.

<sup>27</sup> Transcript p. 59.

<sup>28</sup> Transcript p. 48.

<sup>29</sup> Transcript pp. 60-1.

<sup>30</sup> Transcript p. 60.

Education’s (MSDE) state test security officer coordinates MSDE staff members who make “random observations at schools across the state to observe testing to ensure that standardized testing procedures are followed and to obtain feedback for improvement of future testing.” This is done for each of Maryland’s assessments. All MSDE staff who participate in monitoring must complete a training session facilitated by the State Test Security Officer that addresses “what Monitors should do before, during, and after observations, procedures for scheduling observations and for visiting the schools, and what to look for during an observation.” Within two days of making an observation, a monitor must complete an observation form and submit it to the project manager for the assessment as well as to the State Test Security Officer. Staff members review the forms “to determine whether there are any test security concerns that must be addressed and to note any concerns regarding the administration.”<sup>31</sup>

Louisiana State Department of Education staff who conduct monitoring visits are trained and visit “anybody who has [had] a prior problem [with testing irregularities]. [State Department staff] single those out.” Additional monitoring visits occur at random. The SEA announces that it will try to visit each school district, “just so they know we’re coming,” but does not identify which specific schools in the districts will be monitored.<sup>32</sup> Several respondents to the RFI also indicated that unannounced monitoring visits play an important role in preventing testing irregularities.

## **Section II. Detection and Analysis of Irregularities in Academic Testing**

This section describes how school officials can use different analyses to detect testing irregularities, how the results of these analyses ought to be interpreted, and whether and how different types of analyses can be used to complement one another. The discussion considers the uses and limitations of these methods.

### **Build irregularity detection into the testing process.**

Ms. Whitehead emphasized that a test security plan must be built into the test administration process that is capable of detecting irregularities that occur before, during, and after the test administration. A strong test security plan should include an audit process to detect tampering by students and professionals before the test is administered, and a post-analysis to detect irregularities such as widespread evidence of unusual erasures.<sup>33</sup>

According to Ms. Fincher, a test security plan should also address the information that should be collected at the school and classroom level to support data forensics and investigations. For instance, certain information – such as proctor and class assignments and seating charts – should be recorded on the day of the assessment’s administration because it can be difficult to collect later.<sup>34</sup>

### **Begin monitoring for irregularities prior to test administration.**

---

<sup>31</sup> Tamara Lewis on behalf of Maryland State Department of Education, General Methods By Which Potential Test Security Threats Are Identified, p. 2, February 2012.

<sup>32</sup> Transcript p. 62.

<sup>33</sup> Transcript p. 106-9.

<sup>34</sup> Transcript p. 124.

The most common cause of irregularities prior to testing is the compromise of test materials. These irregularities can have far-reaching consequences, making the early detection of test security breaches essential to preventing widespread compromise.<sup>35</sup> Test security plans should include an audit that can detect tampering that occurs before the test is administered. In an effective audit, test materials are individually inventoried to inspect for missing test booklets or signs that tests or items have been reproduced. School officials should be notified immediately if any irregularities are found. Generally, the most effective audits are unannounced.<sup>36</sup>

**Proctors should monitor for irregularities during the test administration.**

Active proctoring is essential for identifying testing irregularities during administration.<sup>37</sup> Proctors should check student identification to protect against impersonation. “Wandering eyes” during the administration can indicate copying or other forms of communication. Proctors should check for any unauthorized aids, such as cell phones, that could be used to transmit or receive answers or otherwise signal answers during test administration. When collecting answer sheets, the proctor should inspect each one to check for impersonation or substitution of the answer sheet.<sup>38</sup>

**Use comprehensive integrity analyses to identify irregularities post-administration.**

According to NCME’s testing integrity guidelines, SEAs (or their designees) “should conduct comprehensive integrity analyses at multiple levels” (*i.e.*, at the student, class, school, and district levels) for “all large-scale programs where consequences for students and/or school personnel are present.” State analyses and reports “typically provide the best comparison for evaluating schools and districts,” and “should be reviewed by the SEA’s technical advisory panel.”<sup>39</sup>

Analyses should be comprehensive and should include multiple methods to ensure the highest likelihood of detecting misconduct. Data forensics and statistical analysis methods (*e.g.*, ratio analysis/erasure analysis, item-response pattern analysis, and test-score analysis) can be effective tools in identifying suspicious testing patterns. Unlike other detection methods, such as tip lines or studies of individual schools, statistical analysis can analyze “entire populations of classrooms in an area,” measure the scope of an irregularity, and “can identify systematic patterns as well as individual cases of concern.” Further, statistical analysis is cost-efficient. It costs relatively little and can cover large testing populations. According to Professor Jacob, there are three primary approaches for detecting testing irregularities using statistical analysis:<sup>40</sup>

*Ratio analysis/erasure analysis*

All machine-scored, multiple choice tests can be scanned to identify the presence of a second mark on an answer sheet in response to an item. This second mark often indicates that a student erased and changed his or her answer. By evaluating the number of answers that were changed from an incorrect to a correct response for all students in the state, an SEA can use ratio analysis to identify patterns that warrant

---

<sup>35</sup> Transcript p. 106.

<sup>36</sup> Transcript p. 107.

<sup>37</sup> Transcript pp. 107-8.

<sup>38</sup> Transcript p. 108.

<sup>39</sup> National Council on Measurement in Education (2012). *Testing and Data Integrity in the Administration of Statewide Student Assessment Programs*, p. 5.

<sup>40</sup> Transcript p. 99.

further investigation. An unusually high number of wrong-to-right erasures can be an indicator of testing irregularities.

#### *Item-response pattern analysis*

Item-response analysis examines whether there are unusually common response patterns across students within the same class. When using this type of analysis, it is important to account for student characteristics that could explain any common response patterns. According to Professor Jacob, “[c]ommonality of item and response patterns across students might be due to curricular instructional focus of the teacher. And obviously large test score gains could be the result of an extremely effective teacher.”<sup>41</sup>

#### *Test-score analysis*

Test-score analysis examines test scores to see if there have been unusually large gains from the previous year or large relative declines in the subsequent year. The procedure identifies patterns in responses to test items. This analysis can look at individual student-level data as well as classroom or school-level results to identify unusual trends.<sup>42</sup>

Statistical analysis has significant limitations. According to Professor Jacob statistical analysis can only establish an inference of cheating; it cannot determine conclusively whether cheating has occurred. Some detection methods result in a “high rate of false positives and/or false negatives.” For example, even if a statistical analysis yields unusual erasure patterns, such patterns are not necessarily the result of cheating. Unusual erasure patterns could be due to students employing legitimate test-taking strategies that result in frequent wrong-to-right erasures. Further, because statistical analyses are inferential, it is “impossible,” according to Professor Jacob, for these methods to identify the individuals responsible for any alleged testing irregularity. SEAs should combine multiple analyses to identify schools that require additional investigation.<sup>43</sup>

### **Section III. Response and Investigation of Alleged and/or Actual Testing Irregularities**

This section addresses the response and investigation of alleged testing irregularities. Panelists and some RFI respondents provided input on the appropriate roles for parties involved in the investigation of alleged testing irregularities. They analyzed strategies for effectively responding to allegations, including forensic analyses, as well as cooperative practices between SEAs and LEAs and barriers to investigating irregularities. Finally, they provided recommendations for managing wrongdoing, including the appropriate imposition of strict and meaningful sanctions; standards of professional conduct, laws, or regulations that dictate the type of sanctions imposed; how the intent of wrongdoing should be determined and by what entity; and how to restore the credibility of a school system in which wrongdoing has been alleged or proven.

---

<sup>41</sup> Transcript p. 103.

<sup>42</sup> Transcript p. 100, 103.

<sup>43</sup> Transcript p. 99.

### **Investigation procedures are in their “infancy.”**

According to Mr. Ferrara, existing procedures for investigating cheating committed by adults on academic assessments are still in their “infancy” or “have not even gotten to birth yet.”<sup>44</sup> Mr. Ferrara noted that a lot of developmental work needs to be done in this area, and, as a result, there is no common, tested standard that school officials can readily rely upon when investigating allegations of cheating.<sup>45</sup>

### **Focus resources on prevention rather than remediation.**

Limited resources should be directed toward the prevention of cheating rather than investigation and remediation. Responding to allegations of systemic cheating is tremendously costly in terms of money and manpower, and can potentially destroy a school district’s credibility with the public.<sup>46,47</sup> In terms of resources, “[p]revention is cheaper” and school districts should invest their limited resources to establish a testing culture that espouses “play[ing] by the rules,” says Robert Wilson, one of the special investigators appointed by Georgia to investigate irregularities on the CRCT in the Atlanta Public Schools.<sup>48</sup>

### **Establish processes that are conducive to conducting an investigation if one is required.**

Procedures for investigating possible testing irregularities should specify what documentation should be collected during testing, provide adequate protection for those reporting alleged misconduct, and detail reporting systems and requirements.

### *Test administration records and retention*

Several panelists emphasized that SEAs and LEAs must lay the foundation for conducting an investigation even before an irregularity is reported. For example, it is important to have in place clear record retention policies that document critical facts about the test administration (e.g., names of the persons administering the test, time and location of the administration, which students had possession of the test booklets). If an investigation ensues, these records will provide essential facts to the investigative team.<sup>49</sup>

### *Whistleblower protection*

---

<sup>44</sup> Transcript p. 206.

<sup>45</sup> Mr. Wilson stated that “up until fifteen years ago, the tests were all about the students.” As a result, although cheating did occur, it was “almost always” by students cheating, not cheating committed by adults. That “equation” shifted when, according to Mr. Wilson, under No Child Left Behind (P.L. 107-110), the test “became more about the teachers and the schools and the principals than it did about the students.” This shift explains the need for professional investigators in the school context because we are no longer “talking about adults to students,” but about “adults to adults that requires real investigation.”

<sup>46</sup> Transcript p. 167.

<sup>47</sup> Mr. Wilson, who led the investigation of the Atlantic Public Schools (APS) cheating allegations, described the scale and cost of the investigation to State taxpayers. The APS investigation team “included over 60 State agents, attorneys, paralegals,” conducted a total of 2,200 interviews, and reviewed over 800,000 documents in a ten-month span. The investigative team concluded that over 180 teachers and nearly 40 principals were “involved ultimately in the end by name.” Further, the investigation diverted APS from its educational mission, and damaged the District’s credibility. He felt, however, that a transparent and thorough investigation of such allegations can help restore a District’s credibility.

<sup>48</sup> Transcript pp. 208-9.

<sup>49</sup> Transcript pp. 124-5.

It is also essential to establish a culture in which reporting – including the self-reporting of inadvertent errors – is expected and commonplace behavior. According to Mr. Wilson, it is incumbent upon state and school district leadership to foster such a culture by making people feel safe, recognizing the value of an honor system, and encouraging reporting. At a minimum, appropriate policies against punishing whistleblowers must be in place in order to ensure that whistleblowers are not persecuted, “so the information can be forthcoming.”<sup>50</sup>

#### *Multiple reporting systems*

School systems can also encourage reporting of possible testing irregularities by having multiple reporting mechanisms. For instance, in Maryland schools can receive anonymous reports via multiple access points: telephone, e-mail, letter, fax, or an in-person conversation.<sup>51</sup> School systems should have call-in or web-based hotlines so that educators, parents, and students can easily report suspected irregularities. When a tip has been received, systems should ensure that each report is logged and tracked, and that “there’s an accountability system for follow-up.”<sup>52</sup> North Carolina, for instance, employs an online system to record and track reports of irregularities. According to Mr. Fabrizio, “We [in North Carolina] have an actual online system where data is entered online and can be tracked. All the steps of the process are in that system. We require local investigations by the LEA. We have staff whose job it is to assist the school districts if they have any questions about how to conduct [an] investigation, and any serious irregularities that get identified require, then, the use of a checklist that must then be submitted.”<sup>53</sup>

#### *SEA, LEA, and school reporting requirements*

Once a report has been received, Mr. Wilson stated that it ought to be “report[ed] up the line.” If a staff member or teacher receives the report, he or she should report to school leadership – the test coordinator, the assistant principals, or principal. School leadership then should report it to the school district’s leadership – the assistant superintendents or superintendent. School district leadership should then involve legal counsel, and notify the SEA.<sup>54</sup>

#### **Establish a standard or trigger for an investigation.**

All of the panelists agreed that the bar for triggering an investigation should be low. Any allegation or report of cheating – even if just hearsay or gossip – should initiate an investigation. According to Mr. Wilson, school officials should follow up on even gossip and other casual conversation because members of a school community “don’t know they are reporting something, and they actually are, through casual conversations.” For instance, a school official may receive a tip at a meeting when a parent innocently states that “I don’t really understand these test scores because I know that my son or daughter is not performing at these levels.”<sup>55</sup>

Mr. Wilson asserted that follow-up should occur even if the tip or report is vague or difficult to investigate. In such a situation, the school system still has an obligation to exercise due

---

<sup>50</sup> Transcript p. 133.

<sup>51</sup> Tamara Lewis on behalf of Maryland State Department of Education, General Methods by Which Potential Test Security Incidents are Identified, p. 1, RFI Response.

<sup>52</sup> Transcript pp. 197-8.

<sup>53</sup> Transcript p. 185.

<sup>54</sup> Transcript p. 159.

<sup>55</sup> Transcript p. 197.

diligence by speaking with others to determine if the report has any credibility.<sup>56</sup> For example, when North Carolina receives a vague anonymous tip, the SEA immediately calls the central testing office of the school district, and directs the district's testing office to investigate the matter.<sup>57</sup>

**If necessary, trained personnel ought to conduct an investigation.**

When responding to a report trained personnel usually should conduct the investigation rather than school personnel. In a situation where the alleged cheating appears to be a limited and isolated event, it may be appropriate for school personnel to investigate the report; however, if the report suggests cheating on a larger, more systemic scale, it is critical to find trained personnel outside of the school to investigate the matter.<sup>58</sup> Mr. Wilson said that “[s]chool people are not really well-equipped to be investigators. That’s not what you are trying to do -- you are trying to educate. That’s why [you’ve] got to use resources of attorneys, investigators, law enforcement where appropriate... there are a lot of tricks to the trade to proper investigations and that is just not what educators are taught to do.”<sup>59,60</sup> Mr. Ferrara agreed and added that due to a potential conflict of interest, “there are reasons why you don’t want to put a school leader or manager in the position of having to investigate one of their staff.”<sup>61</sup>

An investigation conducted by trained personnel is also more likely to withstand legal challenges that result from the investigation. In fact, LEAs and SEAs “should plan for litigation. There is no way [to] have a strong investigative protocol and process [and] not be involved in future litigation.”<sup>62</sup> Investigations that adhere to clearly delineated principles and procedures are more likely to withstand legal scrutiny.<sup>63</sup>

**SEAs and LEAs should work together during investigations.**

Once an investigation begins, the panelists and respondents to the RFI emphasized that an SEA and an LEA should work together and share resources.<sup>64</sup> Some states require such cooperation by regulation. For instance, in Oregon, the ODE has “established clear requirements to LEAs to investigate and report testing improprieties and has implemented a uniform reporting procedure that all LEAs must follow.” These requirements include having LEAs “submit specified information about each test impropriety, ensuring consistent investigation practices.”<sup>65</sup>

The SEA and the LEA have their own roles to play during an investigation. In Louisiana, the SEA relies on LEAs to take on some of the detection and investigative work because the LEAs “are there when the tests are given” and are “closest to the source of the problem if there is one.”

---

<sup>56</sup> Transcript p. 203.

<sup>57</sup> Transcript p. 202.

<sup>58</sup> This can include school district staff. In its response to the RFI, Oregon noted that LEAs are often in the “best position to conduct interviews with staff and students during the investigation phase and to take follow-up actions in the aftermath of a test impropriety, such as providing refresher training for staff or to implement tighter protocols to ensure security of the test environment.”

<sup>59</sup> Transcript p. 190.

<sup>60</sup> Transcript p. 162.

<sup>61</sup> Transcript p. 190.

<sup>62</sup> Transcript p. 172.

<sup>63</sup> Transcript p. 181.

<sup>64</sup> Transcript p. 210.

<sup>65</sup> Oregon Department of Education Testing Integrity [RFI] Response, p. 5, February 13, 2012.



Although the SEA is “in charge of erasure analyses and plagiarism, the districts can find many or lots of other things” as well. If the LEA confirms that cheating has occurred, the LEA can either report these findings to the SEA or void the scores on its own.<sup>66</sup>

**Interview all witnesses, and re-interview them, if necessary.**

Mr. Wilson recommended that investigators interview all witnesses and re-interview witnesses who may have information, outside of the school environment, if necessary. He emphasized that student interviews should be conducted carefully and in a sensitive manner.<sup>67,68</sup> Mr. Ferrara added that it was important for investigators to be mindful of the issue of false confessions, and that this was an issue that deserved further attention.<sup>69</sup>

**Transparency can help restore credibility in the wake of an investigation.**

During the course of an investigation, all findings should be reported to the proper internal authorities on a regular basis. In Baltimore keeping the CEO informed and being transparent about the investigation’s findings were critical factors in undoing some of the damage to the school district’s reputation and helped to re-establish some credibility.<sup>70</sup>

Ms. Edwards emphasized the importance of a well-conducted investigation as a means of restoring credibility within the school district, and establishing an appropriate tone at the investigation’s outset. The investigation’s mission should not just be about “catching cheaters,” but it should be about “protecting educators, and...protecting the results of children.”<sup>71</sup> In Baltimore, transparency was key -- “we knew that we had to have the public continue to buy into our school system and know that we were going to do the right thing for our children.”<sup>72</sup>

In its response to the RFI, Caveon Consulting Services noted that restoring credibility “is a very difficult job and many actions must be taken if the offense is substantial,” and “getting external help with proper procedures and communications is recommended.” In order to deter future acts of cheating and to provide assurance to the public, Caveon advised that prompt and decisive action is crucial for restoring credibility. Clear statements of intolerance for misconduct must be made and steps taken to back up those statements.<sup>73</sup>

**Use appropriate sanctions.**

If an investigation concludes that cheating has occurred (*i.e.*, the initial report is supported by solid data and evidence<sup>74</sup>), SEAs and LEAs should be clear about what sanctions the SEA may impose. Depending upon the jurisdiction, “[s]anctions can be imposed ranging from reprimands or revocation of license” to more stringent penalties. In some states, “criminal prosecution is possible as well as suits to recover monetary damages due to loss of valuable intellectual

---

<sup>66</sup> Transcript p. 55.

<sup>67</sup> Transcript pp. 161-3, 191.

<sup>68</sup> Transcript p. 162.

<sup>69</sup> Transcript p. 192.

<sup>70</sup> Tamara Lewis on behalf of Maryland State Department of Education, General Methods By Which Potential Test Security Threats Are Identified, p. 1, February 2012.

<sup>71</sup> Transcript p. 209.

<sup>72</sup> Transcript pp. 172-3.

<sup>73</sup> John Fremer for Caveon, Response to US DOE RFI to Gather Technical Expertise Pertaining to Testing Integrity, pp. 15-6.

<sup>74</sup> Transcript p. 171.

property.”<sup>75</sup> Nonetheless, while states are obligated to follow and enforce their test security procedures, they also must provide adequate due process to the person or persons being investigated. Each person under investigation is generally entitled to notice and an opportunity to be heard.

According to NCME’s guidelines, sanctions should be appropriate in light of all of the circumstances of the situation, and “be proportional relative to the offense and equivalent to other policies.” Due process rights and appeal procedures should be established for staff members suspected of misconduct, and the “accused should be informed of the allegations or complaints and the circumstances behind them (statistical detection, reported violation, etc.).”<sup>76</sup>

The Maryland State Department of Education “recommends sanctions based on sanctions issued in similar cases across the state,” but the LEA makes the final determination. In Maryland, sanctions are proportionate to the severity of the offense, and take the form of progressive action “that may begin with a verbal warning, progressing to a written letter of warning, a written letter of reprimand, removal from duties, suspension, termination, and/or loss of credentials.”<sup>77</sup>

Mr. Norton stated that in Louisiana State policy allows the SEA to invalidate test scores if an investigation (with several layers of review) determines that there was misconduct (*e.g.*, plagiarism). This may be done even if the LEA disagrees with the investigation’s findings.

## **Section IV. Testing Integrity Practices and Procedures for Online and Technology-based Assessments**

In this section, participants address testing integrity practices and procedures as they relate to online and technology-based assessments. Because many academic tests are transitioning from traditional paper-and-pencil format to computer-based testing (CBT), panelists focused on how the responses to the questions in the preceding sections would differ when applied to CBT. They identified processes to ensure, to the greatest extent possible, that the results of CBT are accurate and free from tampering. Finally, they discussed likely opportunities for tampering and testing irregularities within the context of computer-based assessments.

### **Transition to CBT.**

The panelists acknowledged the promise of CBT, but expressed concerns regarding the transition from paper-and-pencil tests to this new format.<sup>78</sup> They agreed in general that CBT

---

<sup>75</sup> John Fremer for Caveon, Response to US DOE RFI to Gather Technical Expertise Pertaining to Testing Integrity, p. 15.

<sup>76</sup> National Council on Measurement in Education (2012). Testing and Data Integrity in the Administration of Statewide Student Assessment Programs. P. 6. .

<sup>77</sup>Tamara Lewis on behalf of Maryland State Department of Education, General Methods By Which Potential Test Security Threats Are Identified, p. 2, February 2012.

<sup>78</sup> Similar to paper-based assessments, technology-based assessments come in a wide variety of forms.<sup>78</sup> According to the Association of Test Publishers, technology-based assessments can be delivered using a desktop computer (Computer Based Testing, CBT) or in other forms such as hand-held devices, tablets, or other personal digital assistant devices. Technology-based assessments do not necessarily need to be

would reduce or eliminate some test security threats that were inherent in paper-and-pencil tests. For instance, with CBT it would no longer be necessary to take measures preventing unauthorized access and tampering with test booklets.<sup>79</sup> Still, some test security threats, such as the inappropriate provision of administrative assistance to students before and during the test administration or students accessing outside resources using a variety of technological devices, would continue to exist under CBT.<sup>80</sup>

CBT also creates new threats to testing integrity. As Caveon Consulting Services noted, “[s]hifting to a new assessment delivery model such as computer delivered or even computer adaptive testing does not make cheating and test piracy problems go away. They merely take a different form.”<sup>81</sup> For instance, the delivery of electronic tests to test sites is “not a perfect process” and “some very tech-savvy folks are going to try to undermine [that process].” Mr. Buckley noted that technology provides students and adults with the platform to widely and rapidly disseminate knowledge about test items “everywhere -- everywhere with [Internet] access -- which means chain of custody issues will also be very important and harder to track.”<sup>82</sup>

#### **Build capacity to ensure secure administration of CBT.**

Many of the risks to the integrity of CBT stem from a lack of capacity and infrastructure to administer CBT in a secure manner. For instance, some schools lack sufficient computers, electrical hookups or other capacities needed to administer CBT assessments to all of their students simultaneously. This tends to result in an extended testing window for CBT, which increases the opportunity for students who have taken the test and been exposed to the item pool to disclose these items to students who have not yet taken the test. This issue of an extended testing window “could present the single largest risk to CBT,” according to Wayne Camara, vice president for research and development at the College Board. Schools need a very low ratio of students to computers to mitigate this risk.<sup>83</sup>

#### **Train to ensure secure CBT administration.**

Additional threats to the integrity of CBT include a lack of training and familiarity with the new technologies. Even though “most people are invested in administering the system to the best of their ability,” CBT can be “overwhelming for new teachers and substitutes” and most inadvertent errors occur “due to a lack of training” in CBT, according to Tony Alpert.<sup>84</sup> Panelists advised that states and school districts should prepare administrators with simulated CBT, and provide clear protocols and help-desk support.<sup>85</sup>

---

delivered through an “online” or web-based format (through the Internet or World Wide Web). Assessments can be proctored locally through workstation-based hardware and servers. Local and temporary hard drives may also save assessment scores and records (such as a CD, USB flash drive, or floppy disk).

<sup>79</sup> Transcript p. 234.

<sup>80</sup> Transcript pp. 235-6.

<sup>81</sup> John Fremer for Caveon, Response to US DOE RFI to Gather Technical Expertise Pertaining to Testing Integrity, p. 4.

<sup>82</sup> Transcript p. 273.

<sup>83</sup> Transcript p. 259.

<sup>84</sup> Transcript p. 254.

<sup>85</sup> Transcript pp. 253-4.

Technology provides students and adults with the platform to widely and rapidly disseminate knowledge about test items “everywhere – everywhere with [Internet] access – which means chain of custody issues will also be very important and harder to track.” Because of this, school systems should develop plans to address the chain of custody, monitoring, and other problems posed by new technology.<sup>86</sup> Effective plans should include regular monitoring of the Internet and other media for sharing secure information, routine reporting of monitoring results, and regular auditing of the monitoring program to assess its effectiveness.<sup>87</sup>

**Clearly delineate new processes to mitigate threats posed by the transition to CBT.**

Many of the risks unique to CBT can be mitigated by implementing processes developed explicitly for this mode of assessment. In fact, the “vast majority” of irregularities are mistakes caused by a “lack of clear delineation by the state that’s providing the assessment.”<sup>88</sup> When transitioning to CBT, states and school districts will have to develop policies and internal controls for addressing new kinds of data breaches. CBT security policies will, by necessity, involve a larger network of adults beyond teachers and proctors, such as information technology staff. Policies will need to address logistical challenges such as protocols for logging off or securing computers when students leave to use the restroom or if a computer is intentionally or unintentionally disabled (*e.g.*, unplugged) during a test.<sup>89</sup>

In Oregon test vendors must deliver the state’s online assessment “within a secure browser that restricts students’ access to the desktop and Internet based on requirements set by the ODE.” This secure browser prevents students from accessing other applications including the Internet while testing; it also prevents students from taking screen shots of the online test or copying or pasting data. In addition, “students are required to log in to the secure browser using a secure student identifier and to receive approval from a test administrator before accessing the online test.”<sup>90</sup>

CBT security policies should contain a variety of components, including the following:

- Limiting or disabling web browsers to prevent access to resources on the Internet;<sup>91</sup>
- Limiting or disabling computer applications, such as spell-check, calculators, and other tools;
- Disabling “screenshot” abilities (the ability to take a digital picture of the screen);<sup>92</sup>
- Disabling save, copy, and print functions;<sup>93</sup>
- Conducting CBT only in secure classrooms or computer labs;<sup>94</sup>
- Supervising test access through the use of secure student log-in identifiers or serial numbers;<sup>95</sup> and

---

<sup>86</sup> Transcript p. 273.

<sup>87</sup> John Fremer for Caveon, Response to US DOE RFI to Gather Technical Expertise Pertaining to Testing Integrity, p. 22.

<sup>88</sup> Transcript p. 253.

<sup>89</sup> Transcript pp. 264-9.

<sup>90</sup> Oregon Department of Education Testing Integrity [RFI] Response, p. 4, February 13, 2012.

<sup>91</sup> Transcript p. 231.

<sup>92</sup> Transcript p. 182.

<sup>93</sup> Oregon Department of Education Testing Integrity [RFI] Response, p. 4, February 13, 2012.

<sup>94</sup> Transcript p. 252.

<sup>95</sup> Ranjit Sidhu for American Council on Testing (ACT), Inc., Testing Integrity Response, p. 12, February 16, 2012.

- Ensuring that technology is well-equipped to guard against hacking and other exposure.

### **Use CBT to detect irregularities during test administration.**

Despite the challenges that the transition to CBT presents, all of the panelists agreed that once the format becomes routine, it will provide numerous advantages over traditional paper-and-pencil testing, especially in terms of improved test security measures. Leveraging the advantages that CBT provides will help school officials dramatically improve the security of the testing process. For instance, “down the road, we’ll be able to detect [cheating] while it’s occurring and do something to interfere with the cheating online.” CBT provides officials with a “treasure trove” of test administration data that paper-and-pencil tests cannot provide, such as response time (how long it takes a student to answer a question),<sup>96</sup> the number of wrong-to-right corrections, the order in which questions are answered,<sup>97</sup> and keystroke patterns. Data can then be analyzed to identify patterns and detect anomalies that occurred during the test administration process, which are “real clue[s]” that cheating may have occurred. For example, if a large number of students have short response-time on hard questions and long response-time on easy questions, this pattern may indicate that the students were coached prior to taking the test.

Despite these advantages, Mr. Bruce cautioned that these are still statistical analyses, and therefore subject to the same kinds of limitations as other statistical methods. Even though it may be possible to reduce the number of false positives using CBT, and obtain stronger and/or additional evidence, computer-based security is still inferential and “still proves nothing. It takes [an] investigation to do that.”<sup>98</sup>

## **Closing Comments**

Federal, state, and local education decisions are based, in part, on the results of assessments. In order to ensure that they are making the best decisions, officials need assessment results that are valid, reliable, interpretable, and accurate. Although the vast majority of our nation’s educators would never participate in or excuse cheating, cheating does and will continue to happen. When it occurs, the public loses confidence in the educational system, and serious educational, fiscal, and political consequences can follow.

The education community must invest in test security to ensure that all students have had equal opportunities to demonstrate their knowledge, skills, and abilities. Still, states and districts face the challenge of limited resources. The best and most efficient way to handle testing irregularities is by preventing them from happening in the first place.<sup>99</sup> In comparison to investigations – which are both costly and potentially damaging to the reputation of an SEA or LEA – prevention of testing irregularities is cost-effective. A clear and comprehensive test security policy can prevent many testing irregularities from occurring. Such a policy should include strong and clear language addressing instructions for test administration, secure management of testing materials, protocol for reporting breaches (e.g., anonymous tip hotlines

---

<sup>96</sup> Ibid at 11.

<sup>97</sup> Testing Integrity Practices and Procedures for Online and Computer-based Assessments, PowerPoint presentation, slide 27, available at: <http://ies.ed.gov/whatsnew/conferences/?id=966&cid=2>.

<sup>98</sup> Transcript p. 249.

<sup>99</sup> Transcript pp. 208-9.

and other reporting systems), explanations of analyses used to detect irregularities, and sanctions for misconduct.

Computer-based testing represents a significant change for jurisdictions that employ paper-and-pencil models, and this transition presents new test security challenges.<sup>100</sup> While CBT will eventually reduce or eliminate threats unique to paper-and-pencil tests, such as unauthorized access to test booklets, CBT also introduces new threats. Lack of capacity and poor infrastructure magnify these threats, and SEAs and LEAs will need time to build the capacity to confront them.<sup>101</sup>

SEAs, LEAs, and the testing companies that serve them are invested in developing and sharing best practices for preventing, detecting, and investigating testing irregularities. The Department values the educational community's willingness to meet this challenge head-on with hard work, fresh thinking, and open, honest, and respectful dialogue, and remains committed to helping educators use reliable data to ensure equal access to education and promote educational excellence throughout the nation.

---

<sup>100</sup> Transcript pp. 232-6.

<sup>101</sup> Transcript p. 259.

## References Cited

Amrein-Beardsley, Audrey, Berliner, David C., Rideau, Sharon, (2010). *Cheating in the first, second, and third degree: Educators' responses to high-stakes testing*. Education Policy Analysis, Vol. 18, Issue 14, pages 1-36.

Maryland State Department of Education (2012). General Methods by Which Potential Test Security Threats are Identified. Unpublished document.

National Council on Measurement in Education (2012). *Testing and Data Integrity in the Administration of Statewide Student Assessment Programs*. Retrieved January 5, 2013 from [http://ncme.org/default/assets/File/Committee%20Docs/Test%20Score%20Integrity/Test%20Integrity-NCME%20Endorsed%20\(2012%20FINAL\).pdf](http://ncme.org/default/assets/File/Committee%20Docs/Test%20Score%20Integrity/Test%20Integrity-NCME%20Endorsed%20(2012%20FINAL).pdf)

U.S. Department of Education (2012). *Request for Information to Gather Technical Expertise Pertaining to Testing Integrity*. Federal Register, Vol. 77, No. 10. January 17, 2012. Retrieved September 30, 2012 from <http://www.gpo.gov/fdsys/pkg/FR-2012-01-17/pdf/2012-753.pdf>

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics (2012). *Transcript of Proceedings. Testing Integrity Symposium. February 28, 2012*. Retrieved September 30 from [http://nces.ed.gov/whatsnew/conferences/PDF/2012\\_symposium\\_transcript.pdf](http://nces.ed.gov/whatsnew/conferences/PDF/2012_symposium_transcript.pdf)

## Appendix A: Testing Integrity Symposium Panelist Biographies

**Tony Alpert** serves as the chief operating officer for the Smarter Balanced Assessment Consortium, where he oversees the fiscal operations of Smarter Balanced, collaborating with the State of Washington as the lead fiscal state; and provides expert guidance on assessment design and technology. Prior to joining Smarter Balanced, Alpert served as the director of assessment for the Oregon Department of Education where he directed the administration of Oregon's computer adaptive assessment system. He also served on the U.S. Department of Education's National Technical Advisory Committee where he provided technical counsel. Alpert earned his master's degree at the University of Oregon.

**Audrey Amrein-Beardsley** is currently an Associate Professor in the Mary Lou Fulton Teachers College at Arizona State University. Her research interests include educational policy, educational measurement, and research methods. Specifically, she researches high-stakes tests and their intended and unintended effects, looking at whether high-stakes tests increase student learning and achievement, while examining the unintended effects of high-stakes testing (e.g., teaching to the test, narrowing of the curriculum, cheating, etc.) that also sometimes result. She has published numerous scholarly articles on these topics.

**Wes Bruce** was named Chief Assessment Officer for the Indiana Department of Education (IDOE) in January of 2009. Wes began his work with IDOE in 1999 as director of the Division of School Assessment and prior to taking his current position he held a number of wide ranging positions within IDOE. Wes has served on a variety of national panels and has made numerous state, regional and national presentations on assessment, accountability and student data systems. He was appointed by U.S. Secretary of Education Margaret Spellings to the first National Technical Advisory Council. He currently serves on the multi-state Partnership for Assessment of Readiness for College and Careers (PARCC) Leadership Team and chairs PARCC's Technology Operational Working Group. He holds an undergraduate degree in psychology from Rice University, Houston, Texas, and a graduate degree in computer science from the University of Charleston, Charleston, West Virginia.

**Wayne Camara** is vice president for Research & Development at the College Board, responsible for managing research and assessment development for programs including the Scholastic Aptitude Tests and Advanced Placement. A fellow of the American Psychological Association (APA), the Association for Psychological Science, the American Educational Research Association (AERA) and the Society for Industrial & Organizational Psychology, he is the immediate past president of the National Council on Measurement in Education and president-elect of AERA Division D (Measurement and Research Methodology). Wayne is also past president of APA's Division of Evaluation, Measurement and Statistics, past chair of the Association of Test Publishers, and has served as an associate editor of the editorial board of journals in education and industrial psychology. He has served on technical groups including the Department of Defense's Armed Services Vocational Aptitude Battery Committee, the National Collegiate Athletic Association research committee, Achieve, and the U.S. Department of Education's National Technical Advisory Committee. He currently serves in technical advisory panels for four states, the American Institute of Certified Public Accountants exam, the Partnership for Assessment of Readiness for College and Careers Common Core assessment consortium, the University of Southern California's Center for Enrollment Research and personnel and licensing testing programs in industry. His research and publications focus on college admissions, college readiness, large-scale assessment, ethical and professional issues in assessment and test validation. Wayne received his Ph.D. in



psychology from the University of Illinois at Urbana-Champaign.

**Gregory J. Cizek** is Professor of Educational Measurement and Evaluation at the University of North Carolina-Chapel Hill. He teaches courses in psychometrics, assessment, statistics, research methods, and program evaluation. His interests include standard setting, validity, test security, and testing policy. He provides expert consultation at the state and national level on testing programs and policy, including service as a member of the National Assessment Governing Board which oversees the National Assessment of Educational Progress. He has served in leadership positions in the American Educational Research Association; he is currently President of the National Council on Measurement in Education. Greg has managed national licensure and certification testing programs and worked on test development for a statewide testing program. He has served as an elected member of a local board of education; he began his career as an elementary school teacher.

**Tisha S. Edwards** is Chief of Staff of Baltimore City Public Schools and founder of Baltimore Freedom Academy, a charter high school in the city that was named one of the nation's best high schools by *U.S. News & World Report*. She attended the University of Maryland and obtained a master's degree in social work and a law degree from the University of Maryland School of Law.

**Lou Fabrizio** is the Director of the Division of Data, Research and Federal Policy at the North Carolina Department of Public Instruction. He is responsible for several federal reports including the Education Data Exchange Network (EDEN) and the Consolidated State Performance Report (CSPR), management of the NC CEDARS longitudinal database and federal policy development. He continues to serve as the state's Federal Liaison with the U.S. Department of Education and also serves as a senior advisor to staff in the Accountability Services Division (where he had worked for 15 years) regarding the state's ABCs Accountability Program, the statewide testing program, and the state's accountability and assessment plans for the Elementary and Secondary Education Act. Lou previously worked in a major test publishing company and has been a Head Start Director and math/science teacher. He has a B.S. in Physics from Georgetown University, and both a M.S. in Education Administration and Supervision and Ph.D. in Education Research and Policy Analysis from North Carolina State University.

**Steve Ferrara** is a Vice President and Co-Director for Performance Assessment at Pearson. Prior to joining Pearson, Steve was a Principal Research Scientist at CTB/McGraw-Hill, Managing Research Director in the assessment program at American Institutes for Research, and State Assessment Director in Maryland. Steve conducts psychometric research and designs large scale assessments for K-12 educational achievement, special education, and English language proficiency assessment programs and has consulted on published research on test security protection and violation detection. Steve's work on test security protection and violation detection began when he directed the Maryland statewide assessment programs (1991-1997) and has continued as he worked on assessment program contracts over the last 15 years. Steve earned a Ph.D. in Educational Psychology (concentration in educational measurement) from Stanford University (1989). He also earned an Ed.S. in Program Evaluation (Stanford University, 1984), a M.Ed. in Special Education (mild and moderate disabilities, diagnostic assessment; Boston State College, 1978), and a B.A. in English (minor in Journalism; University of Massachusetts Amherst, 1973).

**Melissa Fincher** has worked in the field of assessment for more than 20 years. Her background includes experience with assessments in both university and K-12 settings. She has developed and managed a number of large-scale, high-stakes assessment programs for both Tennessee and Georgia. In her present role as the Associate Superintendent for Assessment and Accountability at the Georgia

Department of Education, she oversees the development of all Georgia assessment and accountability programs. She is responsible for ensuring these programs meet high standards for technical defensibility. Melissa has served as the principal investigator in several grants and research projects and has presented at numerous professional conferences.

**David Foster** currently serves as the CEO and President of Caveon, LLC. He is also Chief Scientist and Executive Vice President of Kryterion, a unique Internet test administration company he founded. He directed the certification test development efforts at Novell from 1990 to 1997, introducing innovations with regard to such areas as adaptive testing, testing in multiple languages, and simulations-based testing. David has been President and Chairman of the Association of Test Publishers and served on the Board of Directors for the American National Standards Institute. He founded the Performance Testing Council to further the use of tests that directly measure important job or educational skills. He currently sits on the Council for the International Test Commission. He has authored numerous articles for industry trade journals and textbooks, and has presented extensively at industry conferences. David graduated from Brigham Young University in 1977 with a Ph.D. in Experimental Psychology and completed a Biopsychology post-doctoral fellowship at Florida State University in 1982.

**John Framer** is a founder of Caveon, LLC, a company that works on improving security in test development, test administration, reporting, and score use and which assists some states, school districts, and other entities to investigate testing irregularities. John has more than 45 years of testing experience, including management positions at Educational Testing Service and Pearson. John is a past president of the Association of Test Publishers (ATP) as well as the National Council on Measurement in Education (NCME) and the Association for Assessment in Counseling. John received the 2007 ATP Award for Contributions to Measurement. He served as editor for the NCME journal *Educational Measurement: Issues and Practice*. He is co-editor of the *Handbook of Test Security*. John has a B.A. from Brooklyn College, CUNY, where he graduated Phi Beta Kappa and Magna Cum Laude, and a Ph.D. from Teachers College, Columbia University where his dissertation committee consisted of Walter MacGinitie, Robert L. Thorndike, and Sam Ball.

**Brian Jacob** is the Walter H. Annenberg Professor of Education Policy, Professor of Economics, and Director of the Center on Local, State and Urban Policy (CLOSUP) at the Gerald R. Ford School of Public Policy. He is also a Faculty Research Fellow at the National Bureau of Economic Research and an Executive Committee Member of the National Poverty Center. He has previously served as a policy analyst in the New York City's Mayor's Office and taught middle school in East Harlem. His primary fields of interest are labor economics, program evaluation, and the economics of education. His current research focuses on urban school reform and teacher labor markets. In recent work, he has examined school choice, education accountability programs, housing vouchers, and teacher labor markets.

**James S. Liebman** is the Simon H. Rifkind Professor of Law at Columbia Law School. Jim graduated with a B.A. from Yale in 1974 and a J.D. from Stanford in 1977. He served as law clerk to Judge Carl McGowan of the U.S. Court of Appeals for the District of Columbia Circuit from 1977 to 1978, and to Justice John Paul Stevens of the U.S. Supreme Court from 1978 to 1979. He was Assistant Counsel at the NAACP Legal Defense and Educational Fund from 1979 to 1985 and has been a member of the Columbia Law School faculty since then. In 2006-2009, he took a partial leave to serve as Chief Accountability Officer at the New York City Department of Education. While there, he created and led the Division of Accountability and Achievement Resources, directed the City's assessment programs, led the design and implementation of the City's comprehensive new accountability system, facilitated the establishment in each of the City's 1,500 schools of inquiry teams for using data to improve student learning, and

designed and deployed an award-winning citywide instructional data system. Jim currently directs the Center on Public Research and Leadership, a collaboration of the Columbia Business and Law Schools and Teachers College that immerses graduate and professional students in educational policy, research and consulting.

**Scott Norton** is Strategic Initiative Director for Standards, Assessment, and Accountability at the Council of Chief State School Officers. In this role, he works with states to implement the Common Core State Standards and assessments and to create and implement new student-focused accountability systems. The team is also responsible for the State Collaboratives on Assessment and Student Standards, the National Conference on Student Assessment, and collaborative work with the assessment consortia. He previously served as the Assistant Superintendent of the Office of Standards, Assessments, and Accountability at the Louisiana Department of Education. His responsibilities included the implementation of content standards and development of the Louisiana Comprehensive Curriculum as well as the state's transition to full implementation of the Common Core State Standards. He holds a Ph.D. in Educational Administration and Supervision from Louisiana State University.

**Carswell Whitehead** is a Director in the Educational Testing Service's (ETS's) Office of Testing Integrity, where she plans, directs and manages the security investigations and policies designed to prevent and detect score irregularities for College Board testing programs. She has worked at ETS for 22 years, and has 20 years of experience in the test security field. Carswell received a bachelor's degree from Trenton State College.

**Robert E. Wilson (Bob)**, an attorney for 37 years, is a former three-term (1980-1992) District Attorney for DeKalb County, Georgia. Prior to his election, he served as an Assistant District Attorney and Chief Public Defender. He has worked with every Georgia governor for the past 30 years, with experience that includes conducting four special investigations and serving as Presiding Officer for the State Board of Public Safety. Bob has been in the private practice of law since 1992, and is a principal in the firm of Wilson, Morton & Downs, LLC. His practice areas include general litigation, education and municipal law, and special investigations. He is a graduate of the University of North Carolina at Chapel Hill, and holds a law degree from Emory University School of Law. Bob is one of the Special Investigators appointed by former Georgia Governor Sonny Perdue and Georgia Governor Nathan Deal to investigate cheating on the CRCT in the Atlanta Public Schools and Dougherty County School System. Over the past 25 years, Bob has been named by Georgia Trend magazine as one of the 100 Most Influential Georgians, most recently for 2012.

## Appendix B: Request for Information (RFI) Responses

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff, National Bureau of Economic Research. "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood." December 2011. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0016>.

DeVito, Pasquale J. and Michael Nering, Measured Progress. "Comments on Minimizing Testing Irregularities in Large Scale Assessment Programs." February 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0009>.

Diskey, Jay, on behalf of the Association of American Publishers. "ED-2012-OESE-0003 Testing Integrity Response." February 16, 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0014>.

Fields, Bamby. "Response to Request for Information to Gather Technical Expertise Pertaining to Testing Integrity." February 2, 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0003>.

Fremer, John, on behalf of Caveon Test Security. "Caveon Response to US DOE RFI to Gather Technical Expertise Pertaining to Testing Integrity." February 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0015>.

Haygood, Tamara. "Response to Request for Information to Gather Technical Expertise Pertaining to Testing Integrity." February 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0002>

Kaufmann, Jeremy. "CSPC Fellowship Paper." February 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0011>.

Lewis, Tamara, on behalf of the Maryland State Department of Education. "General Methods by which Potential Test Security Incidents are Identified." February 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0007>.

National Council on Measurement in Education (2012). Testing and Data Integrity in the Assessment of Student Achievement. Unpublished working paper. (Note: the October 2012 publication, Testing and Data Integrity in the Administration of Statewide Student Assessment Programs was used in preparing this report.)

Oregon Department of Education. "Testing Integrity Response (Docket ID: ED-2012-OESE-0003)." February 13, 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0005>.

Richardson, Tonja. "Response to Request for Information to Gather Technical Expertise Pertaining to Testing Integrity." February 5, 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0004>

Schaeffer, Robert, on behalf of the National Center for Fair & Open Testing. "Response to U.S. Department of Education Request for Information to Gather Technical Expertise Pertaining to Testing

Integrity.” February 16, 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0012>.

Shutters, Cecila, Jenna Howard, Susan Ingber, and Don Lambing, Georgia Institute of Technology School of Public Policy. “Response to Request for Information to Gather Technical Expertise Pertaining to Testing Integrity.” February 16, 2012.

Sidhu, Ranjit S., on behalf of the Office of Strategic Initiatives, ACT, Inc. “ACT’s Response to the U.S. Department of Education’s Test Integrity RFI.” February 16, 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0010>.

Stetter, Michael, on behalf of the Delaware Department of Education. “Response to Request for Information To Gather Technical Expertise Pertaining to Testing Integrity.” February 14, 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0006>.

Thieman, Alan J., on behalf of the Association of Test Publishers. “Response to Testing Integrity RFI.” February 16, 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0013>.

Vanlandingham, April. “Response to Testing Integrity RFI.” February 16, 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0008>.

Wever, Dan. “Response to Testing Integrity RFI.” February 2012.

Young, Christopher, on behalf of Wonderlic, Inc. “Response to Request for Information to Gather Technical Expertise Pertaining to Testing Integrity.” February 16, 2012. <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0003-0017>.