The 2015 Conference on Test Security

Fourth Annual
November 4-6, 2015
Lawrence, KS

Hosts

Sponsors

Table of Contents

Welcome

It gives me great pleasure to welcome you to Lawrence, Kansas, and the fourth-annual Conference on Test Security.

The University of Kansas has a long history in educational testing. In fact, this year marks the 100th anniversary of the first use of selected-response questions in a large-scale standardized assessment, the Kansas Silent Reading Test of 1915. This test was developed by Frederick Kelly, a student of E.L. Thorndike and the third dean of the University of Kansas School of Education.

In the past century, we have built on that early work, and through the Center for Educational Testing & Evaluation – one of four centers within the Achievement & Assessment Institute – we design and implement testing programs in our home state of Kansas and 17 other U.S. states.

We hope and expect you all will engage in stimulating exchanges of ideas on topics of great significance in the field. I wish you all a fruitful and enriching time of learning, sharing, and networking, and hope you enjoy the conference and your time in Lawrence.

NEAL KINGSTON
Director, Achievement & Assessment Institute
The University of Kansas

Conference History

The Conference on Test Security began in 2012 as the Conference on the Statistical Detection of Test Fraud and focused primarily on statistical methods. In 2013, an Executive Board was selected to provide guidance. In 2014, the Conference Executive Committee expanded the scope of the conference to include a broader range of test-security subjects, and the name changed to its current form to reflect the broader range of topics and experts presenting at the conference.

General Conference Information

Wifi
Guests will want to access the network "Holiday Inn Lawrence". After connecting to the network, simply open a web browser and accept the terms and conditions. Connection will last for 24 hours after which the process will need to be repeated.

Name Badges
Name badges are not only a distinctive fashion statement, they are an important way for conference staff to identify conference attendees. Please wear your name badge in all sessions.

Conference URL
cete.ku.edu/2015-conference-test-secuity

Local Time
Kansas time zone is: Central Time Zone UTC -6:00

Thursday Night: Downtown Lawrence
Sometimes the best kind of breakout session is the one you didn't have to plan. After getting down to conference business Thursday, you'll have the opportunity to hop on a shuttle bus and experience an evening in our historic downtown, a lively, easily walkable district packed with food, fun, shopping, visual art, live music, and more.

Shuttle Service
Thursday-evening shuttle service from the Holidome to Downtown Lawrence will begin at 5:15 PM and continue through 10:15 PM. Downtown drop-off/pickup will be at two locations: At 10th and New Hampshire streets, and across from the Lawrence Public Library in the 700 block of Vermont Street. The drive downtown is about 15 minutes one way; the loop from the Holidome to the two Downtown Lawrence drop-off/pickup points and back takes approximately 35 minutes to complete.

Learn more about Lawrence
»Local food, arts & entertainment: LAWRENCE.COM
»Visitors guide: EXPLORELAWRENCE.COM
»Local news & weather: LJWORLD.COM

Schedule At-A-Glance

Wednesday, November 4

5:00 pm	Conference check-in
6:00 pm	Welcome reception with cash bar

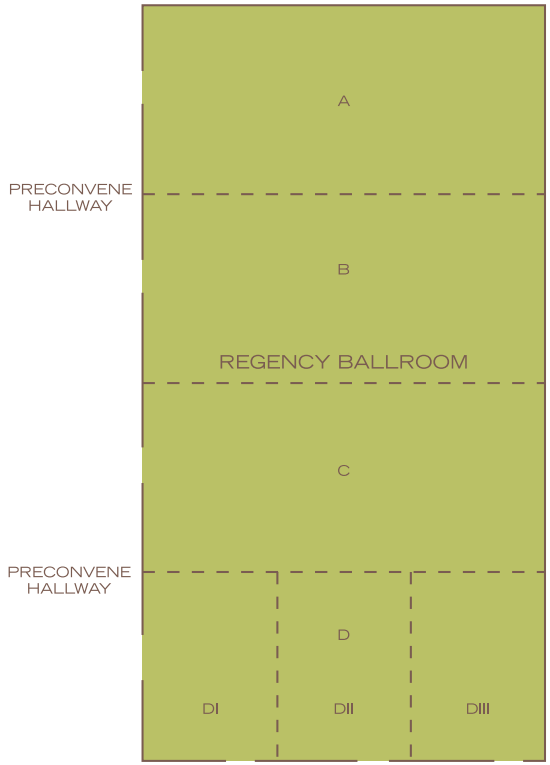
Thursday, November 5

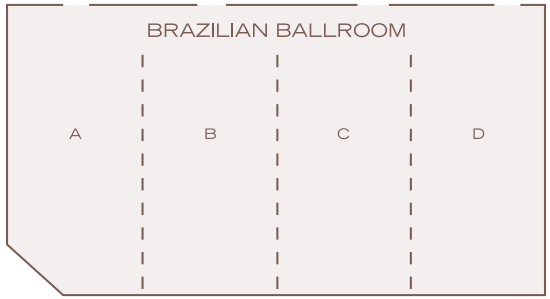
7:00 am	Conference check-in
7:30 am	Breakfast
8:30 am	Keynote
9:45 am	Sessions one
10:45 am	Break
11:15 am	Sessions two
12:15 pm	Lunch
1:15 pm	Sessions three
2:30 pm	Sessions four
3:30 pm	Break
4:00 pm	Sessions five
5:00 pm	Evening off – Shuttles downtown available

Friday, November 6

7:30 am	Breakfast
8:30 am	Sessions one
9:45 am	Sessions two
10:45 am	Break
11:15 am	Sessions three
12:15 pm	Lunch
1:15 pm	Plenary session

Conference Layout





Keynote Speaker

Alan Judd, *The Atlanta Journal-Constitution*

THE TYRANNY OF METRICS: FIELD NOTES FROM A CHEATING SCANDAL
@AlanJudd3000

The widespread cheating scandal that tarnished the image of the Atlanta public school system shows that security measures for standardized testing may be futile as long as such examinations are used as measurements to judge individual performance by teachers and school administrators. The same "data-driven" approach that the Atlanta schools took for a decade has caught on in many fields, from medicine to law enforcement to banking to journalism. In each area, the incessant numerical evaluations can lead to wrongdoing as serious as the school cheating. Ironically, the reporters who uncovered the Atlanta cheating were themselves under pressure to meet performance quotas – and were able to do their job only by ignoring the quotas. Is any security protocol sufficient to overcome the human impulse to meet arbitrary standards by any means necessary?

Alan Judd is an investigative reporter for The Atlanta Journal-Constitution, where he has worked since 1999. Previously he worked for the New York Times Regional Newspaper Group's state capital bureau in Tallahassee, Florida, as well as the Sarasota Herald-Tribune in Florida and The Courier-Journal in Louisville, Kentucky.

Judd was a member of the reporting team that uncovered cheating on standardized tests in the Atlanta public school system, the largest scandal of its kind in the United States. The articles received the Hillman Award for Journalism in 2012 and were a finalist in 2013 for the Goldsmith Award for Investigative Journalism at Harvard University's Kennedy School of Government.

He also has written about suspicious deaths in Georgia's state psychiatric hospitals, which led to a U.S. Justice Department investigation and a long-term plan to reform the facilities. His series on systemic flaws in Georgia's child-protection system led to legislation requiring more transparency in state investigations of children's deaths from abuse or neglect.

Judd is a native of Greensburg, Kentucky, and graduated with a Bachelor of Arts degree in journalism from Western Kentucky University. He lives in Marietta, Georgia, with his wife and their two sons.

Full Conference Schedule

Wednesday, November 4

Conference check-in

Conference desk, outside Regency C
5:00 - 6:00 pm

Welcome reception with cash bar

Regency C
6:00 – 8:00 pm

Thursday, November 5

Conference check-in

Conference desk, outside Regency C
7:00 - 8:30 am

Breakfast

Preconvene Hallway
7:30 – 8:25 am

Keynote

Regency ABC
8:30 – 9:30 am
The Tyranny of Metrics: Field Notes from a Cheating Scandal
» Alan Judd

▷ Session One — 9:45–10:45 am
Statistical Detection I (Papers)

Brazilian A

The Bayesian Flip: An Approach for Quantifying Uncertainty in Probabilistic Judgments of Cheating Evidence
» William Skorupski & Howard Wainer

Robust Statistical Analysis of Answer Changes and Response Times in CBT
» Stephen Cubbellotti & Dmitry Belov

A Comparison of Correlational, Cluster, and Item Response Theory Analyses in the Detection of Compromised Items
» Sarah Thomas, Kim Brunnert, Joy Matthews-Lopez, & Karen Schmidt

Standards and Guidelines (Papers)
Brazilian B

Ensuring Academic Integrity With Online Proctoring
» Zacch Becker

Two Birds, One Clone: Development Considerations for Preventing and Detecting Cheating
» Tara Williams & Jenifer Mutchie

Presentation 1
Brazilian C

An Overview of "Proctoring Best Practices"
» James Wollack, Rory McCorkle, Rachel Watkins Schoenig, & Joe Brutsche

Presentation 2
Brazilian D

Protect, Detect, Respond, Improve – A Holistic Approach to the Test-Security Process
» Marc Weinstein, Benjamin Mannes, & Walt Drane

Coffee Break
Preconvene Hallway
10:45 – 11:15 am

▷ Session Two — 11:15–12:15 pm
Statistical Detection II (Papers)
Brazilian A

Detecting Item Compromise Using Odds Ratio Statistics
» Carol Eckerly, Ben Babcock, & James Wollack

Using Flawed Answer Key Analysis to Detect Braindump Users
» Marcus Scott, Chuck Cooper, & Dennis Maines

Detecting Brain Dump Test Fraud Using Support Vector Machines and the Rasch Model
» Sarah Thomas, Karen Schmidt, Kim Brunnert, & Tim von Oertzen

Other Tools I (Papers)
Brazilian B

Balancing Accessibility and Test Security: K-12 Lessons Learned that Help Inform Licensure and Certification-Exam Policies
» Jill Van den Heuvel, Sheryl Lazarus, & Martha Thurlow

Rogue Sellers and the Banding Together of Higher-Education Publishers
» Kim Brunnert & Nicholas Tardif

Presentation 3
Brazilian C

When You're Not in Kansas Anymore! International Testing From a Practical Perspective
» Rachel Watkins Schoenig, Ray Nicosia & Ardeshir Geranpayeh

Lunch
Regency ABC
12:15 – 1:15 pm

▷ Session Three — 1:15–2:15 pm
Other Tools II (Papers)
Brazilian A

Quantifying the Impact of Compromised Items for Decision Making in Test Security
» Xin Liu, Casey Codd, Christine Mills, & Christy Frederes

Utilizing Interactive, Data-Visualization Software for Prioritizing Case Openings
» Brett Chaney & Jay Parchure

Presentation 4
Brazilian B

Essential Tips for Leveraging Your Vendor's Expertise to Strengthen the Security of Statewide Assessments
» Steve Addicott, Walt Drane, Rachel Watkins Schoenig, & Dennis Maynes

Presentation 5
Brazilian C

Proctoring's Brave New World: Technology-Enabled Proctoring Solutions Delivered Across the Globe
» Chris Kolhouse, George Eftang, John Weiner, & Ruben Garcia

Presentation 6
Brazilian D

Applications of Combinatorial Optimization in Test Security
» Dmitry Belov

▷ Session Four — 2:30–3:30 pm

Presentation 7

Brazilian A

1,001 Easy Steps to Test-Security Conversations: Creating a Culture of Consistent Test-Security Messaging Across Your Sales Organization
» Tara Miller, Christy Frederes, & Karen Wood

Presentation 8

Brazilian B

Establishing Test-Integrity Systems in Coordination: A Look at the Relationship Between State and District
» Tonya Mead & Victoria Nomdedeu

Presentation 9

Brazilian C

How to Develop a Test-Security Flowchart: Critical Documentation for Any Testing Program
» Joy L. Matthews-Lopez & Paul E. Jones

Presentation 10

Brazilian D

Shutting the Barn Door After the Horse Has Bolted
» Susan Weaver & Saundra Foderick

Coffee Break

Preconvene Hallway
3:30 – 4:00 pm

▷ Session Five — 4:00–5:00 pm

Presentation 11

Brazilian A

Security in the Context of Technology Enhancements to Medical Board Certification Assessments
» Linda Althouse & David Foster

Presentation 12

Brazilian B

Online Proctoring: What the Security Professional Needs to Know in a Global World
» Kerri Davis, Nyka Corbin, Joe Brutsche, & Bryan Friess

Posters and Demos

Regency A

Students' Response-Change Behaviors in Computer-Based Testing Environment
» Hongling Wang, Chi-Yu Huang, & Deborah Harris

Identifying Response-Copying Between Test Takers and Un-collocated Cheating Collaborators
» Nooree Huh, Yang Lu, & Chi-Yu Huang

SIFT: Software for Investigating Fraud in Testing
» Nathan Thompson

Seating Charts and Timing and Irregularities, Oh My!
» Jennifer Geraets

Pulse of Security – Security Practices in 2015
» Benjamin Hunter, Rory McCorkle, & Chuck Friedman

CESP – Certification for Test-Security Professionals
» Jamie Mulkey

Detection of Test Fraud in China
» Xiang Kong

Dinner on your own

Downtown Lawrence: see page 5 for details

Friday, November 6

Breakfast

Regency ABC
7:30 – 8:25 am

▷ Session One — 8:30–9:30 am

Presentation 13

Brazilian A

Anatomy of an In-School Test-Fraud Investigation
» Rachel Watkins Schoenig, Marc Weinstein, Benjamin Mannes, & Walt Drane

Presentation 14

Brazilian B

Test-Security Smackdown
» Jamie Mulkey, Christy Frederes, & Tara Miller

Presentation 15

Brazilian C

Design Matters in Monte Carlo Investigations of Aberrant Examinee Response Patterns
» Greg Hurtz, Amin Saiar, & John Weiner

Presentation 16

Brazilian D

Lessons Learned in Improving Test Security for State Assessments: Best Practices and Recommendations for the Prevention and Detection of Cheating
» John Olson, John Fremer, Brian Reiter, Kathy Moore, & Marianne Perie

▷ Session Two — 9:45–10:45 am

Other Tools III (Papers)

Brazilian A

Let's Rethink How Technology Can Improve Proctoring
» Nathan Thompson & Keith Morical

Online Identity Management: Authentication vs. Identification
» Zacch Becker

Presentation 17

Brazilian B

Improving Test-Security Policies, Practices, and Procedures for State Assessment Programs
» John Olson, Michelle Croft, Leila Williams, Joslyn Overby, Jason Kolb, & John Fremer

Presentation 18

Brazilian C

The Media is Knock, Knock, Knocking on Your Door. Now What?
» Richelle Gruber, Steve Addicot, Marc Weinstein, Joe Kammel, & Tamara Lewis

Presentation 19

Brazilian D

Herding Cats: How to Keep Track of Your Test-Security Incidents and Report Your Results
» Tara Miller & Michael Clifton

Coffee Break

Preconvene Hallway
10:45 – 11:15 am

▷ Session Three — 11:15–12:15 pm

Presentation 20

Brazilian A

One Size Does Not Fit All: Making Test Security Configurable and Scalable
» Nathan Thompson & Keith Morical

Presentation 21

Brazilian B

Red Rover, Red Rover! Send Cheaters Right Over!
» Erika Johnson & Susan Weaver

Presentation 22

Brazilian C

Developing a Process for Action: A District's Perspective
» Victoria Nomdedeu

Presentation 23

Brazilian D

There Oughta Be a Law!

» John Fremer, Jennifer Semko, Rachel Watkins Schoenig, & Marc Weinstein

Lunch

Regency ABC
12:15 – 1:15 pm

▷ Plenary Session — 1:15–2:30 pm

Regency ABC

Test-Security Salad: A Test-Security Talk Show

» Jamie Mulkey, Walt Drane, Victoria Quinn-Stephens, Jennifer Cunningham, & Neal Kingston

Abstracts

The Bayesian Flip: An Approach for Quantifying Uncertainty in Probabilistic Judgments of Cheating Evidence

William Skorupski & Howard Wainer

The purpose of the paper is to advocate for using Bayesian inference when using statistical methods to detect cheating. A Bayesian approach for calculating the probability of innocence, given the evidence, is presented. This methodology is compared and contrasted with the frequentist approach, calculating the probability of the evidence, given innocence. These analyses are demonstrated using the Bayesian Flip, an application of Bayes' Theorem described in Skorupski & Wainer (2015). The authors demonstrate that the frequentist null hypothesis testing framework may lead to great overconfidence in the innocence or guilt of an accused cheater. The data for the comparisons are the results of many publicly available newspaper articles describing teachers accused of cheating on behalf of their students. Many such cases have led to "guilt in the court of public opinion" (but not necessarily convictions). The basis for these accusations is generally driven by very small p-values used as sufficient evidence, often as a result of an unusually high number of wrong-to-right erasures in a classroom or school. The authors demonstrate that faulty statistical assumptions and statistical reasoning are often (but not always) responsible for these aberrant results.

Robust Statistical Analysis of Answer Changes and Response Times in CBT

Stephen Cubbellotti & Dmitry Belov

The statistical analysis of answer changes (ACs) has proven to be helpful in identifying possible testing irregularities on large-scale assessments and is routinely performed at some testing organizations. However, the reasons behind ACs can be uncertain because of combinations of both technical and human factors. Existing statistics (e.g., number of wrong-to-right ACs) ignore non-aberrant reasons for ACs thereby creating uncertainty, which may result in a large Type I error. In this study, the information about ACs is used only for the partitioning of administered items into two disjoint subtests: items where answers were unchanged and items where ACs occurred. The statistic proposed measures a difference in performance between these subtests, where, in order to avoid the uncertainty, only final responses are used. For each examinee, the difference in performance is computed as a weighted sum of Kullback–Leibler divergence between corresponding posteriors of ability and Kullback–Leibler divergence between corresponding posteriors of speed. Finally, the subtests can be filtered such that the asymptotic distribution of the statistic is chi-square with one degree of freedom. The performance of the new statistic will be studied on simulated and real responses to a high-stakes CBT in comparison with other popular statistics.

A Comparison of Correlational, Cluster, and Item Response Theory Analyses in the Detection of Compromised Items

Sarah Thomas, Kim Brunnert, Joy Matthews-Lopez, & Karen Schmidt

In this study, we compared the accuracy of correlational, cluster, and Item Response Theory (IRT) analyses in detecting items compromised in a brain dump. The data we used were briefly featured at this conference in 2013 during the Potential Test Fraud Detection Challenge and were obtained after a brain dump of a certification exam was discovered, creating a rare situation in which there were known compromised items and known uncompromised items. Ten statistical flags were used to classify items: inter-item correlations, point-biserial correlations, two and three-cluster solutions, Rasch model parameters (item difficulty, infit, and outfit) and post-hoc estimates of item discrimination, lower asymptote, and upper asymptote from IRT's 2PL, 3PL, and 4PL models, respectively. Building on the work presented in 2013, we investigated the overall accuracy of each statistical flag and the accuracy for known compromised and known uncompromised items individually. Our results showed that point-biserial correlation flags had the highest overall accuracy, followed by inter-item correlation and item-difficulty flags. However, the statistical flags seemed to differ in accuracy between known compromised and known uncompromised items. We will discuss the implications of our findings and how these findings might generalize to other contexts.

Ensuring Academic Integrity With Online Proctoring

Zacch Becker

The presentation will demonstrate how educators may prevent cheating, ensure the academic integrity of distance learning programs, and advance policies designed to reduce incidents of dishonesty online using a number of strategies. The presenter will also share industry research and best practices.

Two Birds, One Clone: Development Considerations for Preventing and Detecting Cheating

Tara Williams & Jenifer Mutchie

Cloning is a relatively quick and cost-effective way for many testing organizations to augment security efforts. However, as testing programs rely more on this method, either in anticipation of or in response to cheating or test theft, pertinent questions arise regarding standards of use; that is, how best to implement, develop, and evaluate the clones. Some of the most significant questions include:

- Which types of modifications tend to deter cheating the most?
- What are effective methods for developing clones that detect cheating (i.e., chameleons)?
- How can cloning be used to turn traditional multiple choice item types into their more secure sibling item type, the DOMC?
- What are creative management solutions for seamlessly incorporating cloning into the test development process?

This session will address these questions by using examples and narrative, data-driven argument, and group discussion.

An Overview of "Proctoring Best Practices"

James Wollack, Rory McCorkle, Rachel Watkins Schoenig, & Joe Brutsche

Proctoring Best Practices is a new publication developed jointly by the National College Testing Association (NCTA) and the Association of Test Publishers Security Committee (ATPSC), during a yearlong collaboration.

Compiled by leaders of both organizations and recognized experts in testing and test security, this document provides a balanced and comprehensive perspective on the subject of test proctoring. It captures best practices for test proctoring in paper-based, computer-based, and online testing environments, and should serve as a guide for test sponsors and/or publishers in drafting their test administration policies or guidelines to evaluate their proctoring practices. In establishing best practices, the contributors gave primary consideration to those practices that promoted the security of the test, maintained standardization, and ensured the fair and respectful treatment of all test takers.

In this session, the presenters will discuss the collaboration, the importance of proctoring and the need for developing best practices, and will review the structure of the document and highlight a number of the key issues addressed therein. Join these testing experts as they share the industry's current thinking around proctor responsibilities and best practices, and address how proctor responsibilities compare and differ across deliver methods.

Protect, Detect, Respond, Improve – A Holistic Approach to the Test-Security Process

Marc Weinstein, Benjamin Mannes, & Walt Drane

All sponsors of high-stakes tests fear waking up to media reports about significant test fraud that impacts their programs. Yet all one need do is Google the words "test cheating" to find numerous recent instances where test sponsors uncovered cases of severe fraud perpetrated by examinees and/or test administrators. Despite their best efforts to deter fraud, all programs that administer high-stakes tests remain susceptible to the threat, which is present across all categories of high-stakes testing, including admissions, licensure, certification and statewide educational assessment. Join seasoned test-security professionals as they share their experiences guiding large-scale test programs in the protection of high-stakes tests, detection of testing irregularities, and responding to and investigating test-security incidents.

In this workshop, following a brief Presentationto introduce the key concepts outlined above, the panel will lead participants through small-group activities that include scenarios taken from actual cases of significant test fraud in certification testing, admissions testing, and in-school educational assessment. By working through these scenarios, participants will come to understand the lessons learned by other test programs, and how they can benefit from a holistic approach to the test-security process that includes a constant cycle of protection, detection, response and improvement.

At the end of this session, the participants will be able to:

- Describe how kiosk-based proctoring solutions can be leveraged to help achieve test-security program objectives.

- Compare how proctoring solutions have been successfully implemented in a variety of organizations.
- Gain insight and ideas into how to implement test-security best practices using technology-enabled proctoring solutions.

Detecting Item Compromise Using Odds-Ratio Statistics

Carol Eckerly, Ben Babcock, & James Wollack

The use of odds ratios to detect examinee preknowledge or item compromise has been suggested by McLeod, Lewis, and Thissen (2003), McLeod and Schnipke (2006), and Obregon (2013). This paper builds on this prior research by combining simulations with longitudinal real data analysis from a national medical imaging certification program to study the behavior of odds ratios under different common testing and preknowledge scenarios. We present a practical guide for the use of odds ratio statistics as part of a program's operational item-bank maintenance. Results indicate that the odds ratio does not have a well-defined distribution; hence, practitioners should not use a fixed, pre-determined critical value to flag items that may have been compromised. Instead, odds-ratio statistics are more informative when change across time is analyzed at the item level. Examples of this analysis will be shown using real data. We also discuss strategies to minimize the effect of user misspecification of model inputs on resulting conclusions.

Using Flawed Answer Key Analysis to Detect Brain Dump Users

Marcus Scott, Chuck Cooper, & Dennis Maines

Brain dumps pose serious threats to test validity, regardless of whether they include all or a portion of the item bank. Detecting brain dump users is a difficult problem. If the brain dump has a flawed answer key, differences between the scores from the brain dump key and the scores from the true key can provide evidence of brain dump usage. The research presented in this paper consists of three components: (1) Detection of when and by whom the items were harvested, (2) Evaluation of methods for detecting brain dump users, and (3) Analysis of how many incorrect answers in the flawed key produce reliable means of detecting brain dump users. The detection of the brain dump theft is based on matching the Presentationorder of the items with the order of the items in the brain dump. Detection of brain dump users is based on the disparity in performance between the items with flawed keys and the items with correct keys. Simulations will evaluate the power of the detection methods. Detection methods discussed in this paper will be presented using data from an actual case. The simulations will be informed by these data to guide the research (e.g., select simulation variables and levels).

Detecting Brain Dump Test Fraud Using Support Vector Machines and the Rasch Model

Sarah Thomas, Karen Schmidt, Kim Brunnert, & Tim von Oertzen

In this study, we combined Support Vector Machines (SVMs) from the field of machine learning with the results of a Rasch model analysis to detect items compromised in a brain dump. SVMs classify distinct classes of scores in such a way that the classes are separated by the maximum possible margin on either side. The analysis method in this study represents a novel method of utilizing SVMs, such that the SVM operates on Rasch model estimates, rather than raw item responses. The SVM was trained on a subset of the data and then applied to the remainder of the data. The data represent an international healthcare certification exam that the test publisher discovered was compromised in a brain dump (N = 13,584). We will discuss the results in terms of which Rasch model estimates were most important in item classifications, the overall accuracy of our SVM, and the future of this methodology for classifying examinees.

Balancing Accessibility and Test Security: K-12 Lessons Learned that Help Inform Licensure and Certification-Exam Policies

Jill Van den Heuvel, Sheryl Lazarus, & Martha Thurlow

One vital component of ensuring exam score validity is test security. Another vital component of score validity is the provision of accommodations for individuals who need them. How can the needs for security and candidate accommodations be balanced to ensure that valid score interpretations are possible for all candidates? This paper explores security policies for K-12 educational assessments administered in states and lessons learned about how to meet accessibility needs while minimizing test-security risks. It addresses how those lessons and policies can be adapted by licensure and certification programs to ensure exam score validity for all candidates without jeopardizing exam security.

Topics that will be covered include: adaptive-technology security considerations, considerations when human access assistants provide accommodations, training considerations, and including accommodations in test security and confidentiality agreements. The 2014 Standards for Educational and Psychological Testing (APA, AERA, & NCME) are referenced to help programs bolster their validity argument in relation to combining test security with testing accommodations. The intent of this session is to share information and start a lively discussion about how to provide test-security measures while maintaining accessibility.

Rogue Sellers and the Banding Together of Higher-Education Publishers

Kim Brunnert & Nicholas Tardif

Some of the world's leading higher-education publishers have banded together against unscrupulous online ("rogue") sellers of textbook test banks. Test banks are a resource that can aid instructors, in their item-based work. Instructors use these supplementary materials to create exams and for grading purposes. To preserve their pedagogical value, such supplemental materials are not generally distributed to the public. Rather, they are provided exclusively to instructors on a limited and restricted basis. Rogue sellers often give the appearance that that they are legitimate businesses. Some even claim to receive the test banks directly from the publishers. These rogue sellers prey upon students looking for legitimate assistance studying without advising of the litany of hidden dangers. Of course, the reality is much different. These rogue sellers, which often sell complete digital copies of the test banks are lying to and harming students. Several publishers are working together to minimize the threat and negative impact of rogue sellers to students, faculty and professors,

and to their businesses. This Presentationwill detail the problem, the struggles in this endeavor, and give insight into some possible short- and long-term solutions. Other test-security tools and methods for protecting the validity of testing results and brand integrity

When You're Not in Kansas Anymore! International Testing From a Practical Perspective

Rachel Watkins Schoenig, Ray Nicosia & Ardeshir Geranpayeh

It's a great big world out there, and testing in Korea isn't like testing in Kansas. This session will provide several case studies of instances in which testing has been compromised and suggest tools to better deter or detect attempts to cheat. Armed with this knowledge, you can approach international expansion with a better understanding of the test-security landscape. Join experienced test-security practioners who provide test security on a global basis for academic, workforce, and organizational assessments for an eye-opening discussion of the challenges of international testing.

Quantifying the Impact of Compromised Items for Decision Making in Test Security

Xin Liu, Casey Codd, Christine Mills, & Christy Frederes

Due to increasing test-breach incidents on high-stakes licensure examinations, test companies often face the challenges of adopting proper actions to mitigate the damages. The high cost of test development often poses a tough decision for test companies on whether to totally abandon the contaminated test or to find leeway in mitigating the damages. The knowledge on the possible amount of damages on test validity will benefit the decision-making process. This study aims to demonstrate a simulation and analysis method in quantifying this impact.

It is believed that the impact is compounded with both the ability level of a student and the difficulty level of an item. The compound impact varies with varying levels of contaminations. Thus, the simulation study of the impact involves three related factors of person ability, item difficulty, and contamination severity. The probability of passing the test is examined by different levels of combinations of these factors. In particular, three levels are considered for each factor: ability (low, average, or high), difficulty (easy, medium, or hard), and severity (10%, 25%, or 35%). The probability of passing is analyzed at both individual student and aggregated institution level.

Utilizing Interactive, Data-Visualization Software for Prioritizing Case Openings

Brett Chaney & Jay Parchure

Each year, millions of college applicants take the ACT, across thousands of test centers and multiple test dates. Although only a small fraction of our examinees obtain scores of which we "doubt the validity" thereof, each test date, this leaves our team of investigators with the daunting task of combing through thousands of cases, flagged by several different statistical analyses. We've developed and will demonstrate the use of several interactive, data-visualization dashboards that empower investigators to efficiently explore the data while connecting multiple statistical output files and searching for trends over time among current and historical data, all in a powerful graphical user interface. Moreover,

this approach facilitates the discovery of previously unnoticed patterns across multiple cases, indicative of a more large-scale issue. Prior to developing these tools, there were rows and rows of numbers across multiple spreadsheets to read through, which was highly time consuming, and many patterns went unnoticed. With these tools, we now use a more hands-on approach, diving in and quickly combining the output datasets from a myriad of sources and times, controlling the visual depiction of multiple dimensions and trends, and drilling down to subsets of interest.

Essential Tips for Leveraging Your Vendor's Expertise to Strengthen the Security of Statewide Assessments

Steve Addicott, Walt Drane, Rachel Watkins Schoenig, & Dennis Maynes

Each year since the passage of NCLB in 2001, the stakes associated with statewide assessments have increased, making test-security breaches more likely, more volatile, and more difficult to handle. The evolving landscape presents new challenges for educators, most of whom have received very little training in deterring, detecting, deciding how to handle security incidents. Fortunately, expertise in test security has been growing among vendors who support statewide assessment programs. A panel of representatives from two state departments of education and two vendors with experience in providing test-security services to states and educational agencies will discuss the expertise of vendors can be leveraged to improve test security. For example:

- When is it appropriate to invalidate a student's test score?
- What kinds of data analysis are appropriate for monitoring test security?
- How can vendors assist with investigations into test-security breaches?

Both panelists from the state departments of education have been dealing with these issues. Both panelists from the vendors have experience in helping states deter, detect, and respond to test-security matters. The value of the panel discussion is to draw out relevant and practical information gleaned through experience, not through theory, that can effectively improve statewide assessment programs.

Proctoring's Brave New World: Technology-Enabled Proctoring Solutions Delivered Across the Globe

Chris Kolhouse, George Eftang, John Weiner, & Ruben Garcia

The mere thought of implementing a technology-enabled proctoring solution as part of your test-security program may send you running for the hills. However, with leading-edge technologies, such as testing kiosks, and a growing body of implementation best practices, technology-enabled proctoring may indeed be in the realm of possibility for many testing organizations. Imagine test-takers having the flexibility to complete a rigorous certification test from their laptop at home or from the comfort of their office, while ensuring a robust level of security that a technology-based proctoring solution can provide. This Presentationwill discuss the implementation of kiosk-based proctoring for two different certification-testing organizations and additional technology enabled proctoring solutions. These programs will share their challenges and successes with designing and implementing large-scale, global proctoring implementations and help you envision the possibilities for your own test-security program.

Applications of Combinatorial Optimization in Test Security

Dmitry Belov

Combinatorial optimization (CO) is concerned with searching for an element of a finite set that would optimize (minimize or maximize) a given objective function. This Presentationwill discuss two applications of CO in test security.

In general, item preknowledge is difficult to detect due to three unknowns: (i) unknown subgroups of examinees at (ii) unknown test centers who (iii) had access to unknown subsets of compromised items prior to taking the test. To resolve the issue of multiple unknowns, two CO methods are applied. First, random search detects suspicious test centers and suspicious subgroups of examinees. Second, given suspicious subgroups of examinees, simulated annealing identifies compromised items.The statistical analysis of answer changes (ACs) has uncovered multiple testing irregularities on large-scale assessments. However, existing statistics capitalize on the uncertainty in AC data, which may result in a large Type I error. Without loss of generality, for each examinee, two disjoint subsets of administered items are considered: (1) items with ACs; (2) items without ACs selected by CO methods to minimize distance between corresponding characteristic Curves. A robust statistic measures the difference in performance between these two subsets, where to avoid the uncertainty, only final responses are used.

1,001 Easy Steps to Test-Security Conversations: Creating a Culture of Consistent Test-Security Messaging Across Your Sales Organization

Tara Miller, Christy Frederes, & Karen Wood

As a test-security professional, how many times have you been on a conference call with sales reps or upper management and cringed when you heard them describe your uber-secret test-security process to a client? Meanwhile, the following thoughts swirled around in your head: “When everyone knows about the process, they start working around it to commit misconduct” or “Is that the way they think that process really happens?” or “We stopped doing that process 2 years ago after the international server fiasco.”

In this session, we will discuss the process that was developed to create internal test-security messaging training. We will also discuss how training was made engaging, educational—and dare we say, fun?—for an audience that encompasses more than 200 client-facing employees.

At the end of this session participants will be able to:

- Identify the gaps in the current language of test security within your organization.
- Pinpoint and establish relationships with key client-facing employees who need test-security messaging.
- Provide tools and takeaways that you can use and incorporate into test-security discussions within your own organization.

Establishing Test-Integrity Systems in Coordination: A Look at the Relationship Between State and District

Tonya Mead & Victoria Nomdedeu

The presenters of this session are state- and district-level test-integrity leaders who have extensive experience in the education sector and at multiple levels (state, district, and school) as an educator, a school psychologist, assessment specialists, and investigators of test security.

State Perspective:

(1) To present the behavioral rationalizations for testing improprieties, cheating and misconduct. Once motivations are clearly understood, policies and supportive training can better address the unique challenges of individual schools and districts.

(2) To present a comprehensive system cutting across all functions, sectors, and levels (classroom, school, district and state) for detecting, deterring, investigation and resolving security incidents.

(3) To share concrete examples to demonstrate the ways in which states and districts can work collaboratively together to develop and improve test-security programs (such as stakeholder participation in establishing policy).

District Perspective:

The district is tasked with determining how best to implement a test-integrity process that is supportive of different school settings and sets clear expectations in test administration through strong communication practices (triangulation between schools and the state), the development of transparent reporting systems, maintaining trust between both the state and schools, and taking action consistently.

How to Develop a Test-Security Flowchart: Critical Documentation for Any Testing Program

Joy L. Matthews-Lopez & Paul E. Jones

A test-security flowchart is a visual protocol used to process flagged (potentially anomalous) examinees in a consistent and unbiased manner. This session will outline what a test-security flowchart is, why it is important, and what role one plays in a comprehensive test-security plan. During the session, we will interact with attendees to develop a sample flowchart. Key components of the flowchart will be mapped to criteria outlined in the AERA/APA/NCME Standards for Educational and Psychological Testing and will contribute to a validity argument for score integrity. This non-technical session will be appropriate for all levels of attendees and will cross all business sectors.

Shutting the Barn Door After the Horse Has Bolted

Susan Weaver & Saundra Foderick

Current security solutions often focus on detecting and responding to security breaches. While important, that is only part of a successful security plan. In this session, we will consider the proactive power of exam and item design in preventing, deterring (and yes

detecting) exam fraud. We will discuss design topics such as the use of secure item types, security-enhanced item stems, item rotation, Trojan horses, and strategies for limiting item exposure. A well-designed exam can help keep happy horses home, safe and secure.

Security in the Context of Technology Enhancements to Medical Board Certification Assessments

Linda Althouse & David Foster

In addition to the initial certification examination, the 24 medical boards require doctors to re-certify every few years, asking them to complete a Maintenance of Certification (MOC) examination. Medical boards are committed to continuously reviewing the way they evaluate the competence and excellence of physicians, recognizing that certification and maintenance of certification activities must not only keep up with medical knowledge and best treatment practices, but also with new technologies. Currently different assessment models and enhancements to the examination are being considered and piloted, with the goal of making the examination more relevant and convenient for the doctor. To the external user, these assessment models seem attractive and reasonable. However, to those in the measurement world, they come with a variety of challenges, many being security related. In the context of maintaining the goal of keeping such assessments secure, this session will present proposed assessment models, along with their advantages and disadvantages. Some of the approaches that will be discussed will include online proctoring, no-proctoring with candidate authentication, use of external web resources, and continuous assessment models.

Online Proctoring: What the Security Professional Needs to Know in a Global World

Kerri Davis, Nyka Corbin, Joe Brutsche, & Bryan Fries

Altering your test program administration in any way is a huge endeavor with many implications to consider: test security, data privacy, and candidate behavior should be at the top of the list.

Gain a greater understanding of the security risks and considerations in online proctoring by learning from the experience of two industry leaders servicing very distinct market verticals.

- Microsoft aligns with corporate strategy providing greater reach and choice for its partners and candidates.
- GMAC diversifies its product offering and strategic partnerships to extend its portfolio beyond the core GMAT exam.
- Each of these organizations is applying this emerging delivery model in unique ways by answering similar questions around privacy, candidate experience, and test security.

Students' Response-Change Behaviors in Computer-Based Testing Environments

Hongling Wang, Chi-Yu Huang, & Deborah Harris

In traditional paper-pencil tests (PPT), erasure analysis through mark-intensity analysis for response changes of examinees to multiple-choice items is widely used as statistical evidence to support allegations of test irregularity. The number of items with responses changed and the number of wrong-to-right changes are the commonly used variables for finding students who have unusual response changes. In computer-based tests (CBT), students' response changes can be recorded by their mouse-clicking behaviors (called click data). Not only can the values of the above variables be obtained, more information (e.g., numbers of times that students visited the items, and the time that students spent in each item) can be obtained. In this study, we will explore what information we can obtain from click data. The comparison of students' response-change behaviors from CBT with PPT will also be reviewed, using randomly equivalent samples of examines who tested in each mode. The findings of the study will help determine if the flagging rules used to flag students with unusual response changes in PPT can also be applied to CBT or if they need to be modified, and what other information from click-data can be used to support the allegations of test irregularity.

Identifying Response-Copying Between Test Takers and Un-collocated Cheating Collaborators

Nooree Huh, Yang Lu, & Chi-Yu Huang

As technology evolves, new methods of cheating also evolve that challenge traditional statistical detection methods in identifying cheaters who collaborate in cheating despite being located at different testing locations. The purpose of this study is to explore a procedure to identify examinees who are physically in different testing locations but actually copy from the same source. A simulation study will be conducted which includes various conditions: 1) the number of items in a test, 2) the number of items with responses in a cheat sheet, 3) the number of correct response items in a cheat sheet, 4) the number of cheating examinees, 5) different alpha levels for criterion, and 6) the ability of examinees who used a cheat sheet. The detection procedure consists of two steps: 1) using a modified ω index to identify examinees whose item responses are highly similar to a cheat sheet; and 2) using three statistics (*lz* index, *H T* index, and score estimation) to further investigate examinees whose score patterns do not fit their abilities. The results of this study will show whether the modified ω index and other three statistical methods can efficiently detect examinees who copy responses from the same source.

SIFT: Software for Investigating Fraud in Testing

Nathan Thompson

SIFT is a new software program that has been developed to help testing organizations investigate the possibility of fraud. It provides a user-friendly interface that allows you to select from multiple collusion indices another other analyses, and provides output in Excel for easy viewing and additional work (e.g. sorting). The primary purpose of the program is to bring the psychometrics of data forensics to more practitioners.

The primary portion of the program is the collusion indices, which include: G2; K; K*; PAIR1; PAIR2; Bellezza and Bellezza (observed and random versions); Harpp and Hogan; and Harpp, Hogan, and Jennings. SIFT also provides functionality for response-time analysis, including the Response Time Effort index (Wise & Kong). Of course, a classical item analysis is also performed, as those statistics remain useful, and are sometimes

integrated into collusion indices. Finally, SIFT also crosses these analyses by two levels of nested locations that users are allowed to specify, such as State and District, or Country and City. For example, users might look at the mean Response Time Effort and number of collusion-flagged examinees in each city where your tests are administered, allowing you to flag problematic test centers.

Seating Charts and Timing and Irregularities, Oh My!

Jennifer Geraets

Management of everything involved in the administration of tests, including admission tickets, rosters, seating charts, examinee instructions, timing requirements, irregularity reports and more, can be challenging for both test staff and testing companies. In this session, attendees will learn about a prototype of an application that aims to make the process easier for test staff, deliver a better and more secure testing experience for examinees, and provide ACT with accurate, timely, and actionable information.

Pulse of Security – Security Practices in 2015

Benjamin Hunter, Rory McCorkle, & Chuck Friedman

Security is a critical consideration for any testing organization. This session presents results from the 2015 survey of security practices, conducted by the ATP Security Committee. These results include responses from testing organizations and vendors, show the practices being used by organizations to address security prevention, detection, enforcement, mitigation, and litigation. The presenters will discuss how vendors and testing bodies see the division of security responsibilities, as well as perceptions related to the effectiveness of various security activities. Finally, the presenters will make recommendations for how test sponsors can address these gaps.

CESP – Certification for Test-Security Professionals

Jamie Mulkey

There's a new certification in town, and it's designed just for you, the test-security professional. This is no ordinary certification program, but then again, the test-security expert is no ordinary profession. The Certified Exam Security Professional (CESP) program recently launched the CESP–Generalist exam in Fall 2015. Stop by and learn about the exam and its requirements for certification. You will be inspired by the program's use of the DOMC item type. You will be in awe of its innovative method for developing the exam and measuring exam performance. You will be astounded by the use of the test's simple but very effective online proctoring methodology. Come by for a test drive!

Detection of Test Fraud in China

Xiang Kong

In China, large-scale and high-stakes testing will have a lot of candidates to cheat. Even more serious is the fact that many candidates are using high-tech equipment for cheating; invigilators in the exams find it difficult to find the cheating candidates, and the candidates' scores are very high, causing great harm to the fairness of the examinations. In order to address the rising tide of cheating, I developed the Thunder Cheating Detecting System. This system can be used in examinations after the candidates answer the information for statistical analysis. Using this system to detect cheating of 3 million students annually, we can detect about 3,000 candidates to cheat. This paper will address the use of this system and the application of the large-scale test in China.

Anatomy of an In-School Test-Fraud Investigation

Rachel Watkins Schoenig, Marc Weinstein, Benjamin Mannes, & Walt Drane

When results count, not everyone will play by the rules. But investigating in-school test fraud isn't the same as dealing with adult test-takers. Students, parents, teachers, unions, state agencies, districts and the public all play important roles in how in-school testing occurs, how test fraud is investigated, and how test-fraud cases are resolved. Yet the interests of each of these stakeholders often conflict with one another. And let's not forget the highly politicized climate around anything related to standardized testing these days. Yet despite these challenges, assessment professionals must have the capability to conduct thorough, fair and effective test-security investigations that enable the collection of reliable evidence upon which testing programs can confidently take appropriate action against perpetrators of exam fraud. Join seasoned professionals as they discuss these issues, and leave more prepared to address these concerns in your own state and testing programs. In this workshop, following a brief Presentationto introduce the key concepts outlined above, the panel will lead participants through small-group activities, including scenarios taken from actual cases of significant test-fraud incidents in statewide educational assessment.

Test-Security Smackdown

Jamie Mulkey, Christy Frederes, & Tara Miller

Are you ready for the Test-Security Smackdown? Test your knowledge and gain insights into test-security best practices as we play a little game of who knows more. Competing in teams, test-security scenarios will be presented. The team with the most examples wins the round. If you are looking to take your security program to the next level, this Smackdown is for you!

This session goes beyond basic test-security concepts and discusses application of practical test-security implementation in real world environments. As part of this session, participants will receive practical test-security planning templates that can be easily implemented in your own testing organization.

Will you leave victorious or defeated? Who will be crowned as Test-Security Smackdown champions?

As a result of this session, participants will be able to:

- Describe examples of test-security threats, best practices, and solutions.
- Describe strategies for implementing test-security policies and processes in various testing environments.
- Apply test-security best practices and tools in your own testing organization.

Design Matters in Monte Carlo Investigations of Aberrant-Examinee Response Patterns

Greg Hurtz, Amin Saiar, & John Weiner

Monte Carlo simulation has been used for evaluating the relative merits of different statistical indices for detecting item response patterns associated with test fraud. Simulation is often used to generate "clean" data according to a response model (e.g., Rasch, IRT) and decisions about the parameters of that model, and then decisions about the selection, operational definition, prevalence, and magnitude of aberrant patterns are introduced into the into the simulated responses. Through these design decisions for a Monte Carlo experiment, strong internal validity can be achieved but external validity and generalizability of results may be limited. The presenters review some of the varied strategies and decisions found in recent published literature, and demonstrate through replication and extension of others' simulation work that design characteristics matter and should not be addressed lightly. The presenters discuss multiple indices of cheating detection, including HT, J2, J3, kappa, and Zmatch, as well as several operational definitions of cheating. The presenters provide some key steps, guidelines, and a decision tree in the careful design of such studies with the intention of forging the path for a unifying framework for design of Monte Carlo studies in the realm of investigating test security.

Lessons Learned in Improving Test Security for State Assessments: Best Practices and Recommendations for the Prevention and Detection of Cheating

John Olson, John Fremer, Brian Reiter, Kathy Moore, & Marianne Perie

Test security for state assessments has become an increasingly important topic as media across the nation buzz with stories of cheating in schools. In this session, information is provided from important new reports/resources published by CCSSO to help states improve test security and implement best practices to prevent and detect cheating. Promising and effective strategies, practices, and procedures that states are using for prevention and detection will be shared.

In 2012-13, in response to growing attention across the country on improving the security of assessment programs, the TILSA SCASS conducted a special project to assist states in proactively addressing potential security problems, resulting in an important and useful report. However, states wanted more information, and in 2015 a follow-up TILSA project was completed to expand the Guidebook and help states further enhance test security. The latest report provides numerous examples of approaches states have implemented to stop cheating on tests and brings together many best practices and procedures of state staff and vendors.

Presenters will share detailed information on several important themes related to test security:

- Recommended methods/approaches/guidance to improve test security in states.
- Lessons learned and best practices for preventing cheating.
- Methodologies for detecting test irregularities/improprieties (e.g. data forensics).

Let's Rethink How Technology Can Improve Proctoring

Nathan Thompson & Keith Morical

Technology has revolutionized much of assessment. However, a large proportion of proctoring is still done the same way it was 30 years ago. How can we best leverage technology to improve test security by improving the proctoring of an assessment? Much of this discussion revolves around remote proctoring (RP), but there are other aspects. For example, consider a candidate focusing on memorizing 10 items: can this be better addressed by real-time monitoring of irregular response times with RP than by a single in-person proctor on the other side of the room? Or by LOFT/CAT delivery?

This Presentationdiscusses the security risks and validity threats that are intended to be addressed by proctors and how they might be instead addressed by technology in some way. Some of the axes of comparison include:

- Confirming ID of examinee
- Provision of instructions
- Confirmation of clean test area with only allowed materials
- Monitoring of examinee actions during test time
- Maintaining standardized test environment
- Protecting test content
- Monitoring irregular time patterns

In addition, we can consider how we can augment the message of deterrence with tactics like data forensics, strong agreements, possibility of immediate test shutdown, and more secure delivery methods like LOFT.

Online Identity Management: Authentication vs. Identification

Zacch Becker

A February 2014 Education Department IG audit found that federal rules regarding identity verification in distance education programs "do not sufficiently mitigate the risks of fraud, abuse, and noncompliance." The audit highlights the need for new standards and regulation regarding financial aid disbursement in distance education. Key distinctions this Presentationwill address are the differences between identity authentication and verification. Identity verification typically equates to logging in with LMS credentials. This is an insecure form of verification that can be defined as a single factor component that relies on seeing the same person consistently return. Multi-factor authentication aims to verify that the correct individual is participating by using a multifactor approach, requiring something students have, something they are and something they know. This analysis will tie in how, and possible reasons why, fraud is growing in online programs across the United States and what institutions can do to combat it.

Improving Test-Security Policies, Practices, and Procedures for State Assessment Programs

John Olson, Michelle Croft, Leila Williams, Joslyn Overby, Jason Kolb, & John Fremer

Cheating and test piracy pose major threats to the validity of test-score interpretation and the credibility of large-scale assessment programs. This session focuses on a variety of

approaches that states have implemented in recent years to make improvements to test security via use of more-effective policies/practices/procedures for the prevention and detection of cheating. Attendees to this session will receive an up-to-date summary of current developments in test security in state and other large-scale assessment programs.

The first presenter will identify current test-security state-codified laws, statutes, regulations, and policies, highlight examples of exemplary statutory and/or regulatory language, and identify key test-security components that are missing from current laws/policies. Three state representatives will present details of important improvements to policies/practices/procedures for the security of their assessment programs. Among the topics to be addressed are guidance on improving security in states, best practices for preventing cheating in schools, and better methods for identifying and following up on testing incidents and irregularities. Many excellent examples will be provided on promising approaches that can be implemented by states and their vendors. The discussant, an internationally known expert on test security, will provide insight and commentary on the current state of affairs in the K-12 assessment world.

The Media is Knock, Knock, Knocking on Your Door. Now What?

Richelle Gruber, Steve Addicot, Marc Weinstein, Joe Kammel, & Tamara Lewis

One of the most disconcerting thoughts for a high-stakes test administrator is the media appearing on the doorstep, informing that a test-security breach may have occurred. This could happen at any time, anywhere, with no advance warning that there may be a problem with the validity of test results. Unfortunately, this uneasiness is justified. This exact scenario has happened repeatedly all over the country to programs of all sizes.

This session will provide:

- Lessons learned through actual experiences in high stakes testing programs.
- Information on dealing with cheating allegations or potential testing validity concerns.
- Crucial strategy for appropriate interactions with the media in the event of an issue.

When a high-stakes testing organization has a good base relationship with the media, a plan in place for dealing with challenging situations before they arise, and a proactive, rapid response, it sets the stage not for a public perception nightmare, but a chance to exhibit sincere interest in the potential issue, dedication to finding the source of the problem, and serious effort toward solving the problem. This session will provide the tools needed to feel prepared when the media comes knocking.

Herding Cats: How to Keep Track of Your Test-Security Incidents and Report Your Results

Tara Miller & Michael Clifton

Do you regularly investigate and respond to test-security incidents as part of your test-security program? With all the information and emails you gather, how do you keep it all straight? More importantly, how do you report these incident outcomes to upper management in ways they care about, while also letting them know what you do all day? Having a test-security incident-management system in place is crucial for reacting, documenting, and resolving test-security incidents. Otherwise, it can be like herding cats. A test-security incident-management system is a consistent way of managing incidents across your test-security program.

In this interactive session, we will present different perspectives on how to manage this process and allow others to add their experience and input as well. The session will provide techniques for practical application of security-incident tracking, responding, and reporting, and will provide a forum for discussion among participants.

Objectives:

- Create the process flow for receiving and responding to incidents.
- Evaluate documentation options which best suits your program needs for tracking and reporting.
- Categorize and quantify your incidents into effective reporting tools to upper management to inform about your actions and results.

One Size Does Not Fit All: Making Test Security Configurable and Scalable

Nathan Thompson & Keith Morical

Development of an organization's test-security plan involves many choices, an important aspect of which is the test development, publishing, and delivery process. Much of this process is now browser-based for many organizations. While there are risks involved with this approach, it provides much more flexibility and control for organizations, plus additional advantages such as immediate republishing. This is especially useful because different programs/tests within an organization might vary widely. It is therefore ideal to have an assessment platform that maximizes the configurability security.

This Presentationwill provide a model to evaluate security risks, determine relevant tactics, and design your delivery solution by configuring test publishing/delivery option around these tactics to ensure test integrity. Key configurations include:

- Regular browser vs. lockdown browser
- No proctor, webcam proctor, or live proctor
- Login processes such as student codes, proctor codes, and ID verification
- Delivery approach: linear, LOFT, CAT
- Practical constraints like setting delivery windows, time limits, and allowing review
- Complete event tracking during the exam
- Data forensics within the system

In addition, we invite attendees to discuss technological approaches they have taken to addressing test-security risks, and how they fit into the general model.

Red Rover, Red Rover! Send Cheaters Right Over!

Erika Johnson & Susan Weaver

Who is on the security team for your exams? Who are the links in your security chain? And how might you strengthen your chain?

In this session we will discuss how program managers, exam designers, psychometricians, item writers, item reviewers, and exam delivery professionals are currently securing their

exams and how they can strengthen their exam-security roles in the future. It takes a strong chain to stop cheaters when they charge! And we all know which link they "test" first: the weak link.

Developing a Process for Action: A District's Perspective

Victoria Nomdedu

Often, districts are tasked by their state agency to investigate, take personnel action, and/or improve their testing protocols or procedures following a state-led review. But, then what?

We'll look at how a district can monitor incident reports and support state-led reviews, and then we'll discuss how the District of Columbia Public Schools (DCPS) is taking action. Much like a state agency, it's important for a school district to establish a process for reviewing irregularities, determining the cause, and revising procedures or taking personnel action, but determining how those decisions are made requires the establishment of a timely action plan. We'll review why it's important for a district to establish strong procedures to administer a test with integrity by improving training and trust among your school communities, providing a method to report incidents, and communicating clear expectations while ensuring flexibility for multiple school environments.

We'll also dive deep into one way DCPS is taking action by reviewing our establishment of a testing integrity review committee and how they're able to make recommendations for personnel action, swiftly and consistently.

There Oughta Be a Law!

John Fremer, Jennifer Semko, Rachel Watkins Schoenig, & Marc Weinstein

Where might legislations or regulations be introduced that would facilitate the challenge of providing fair and valid test results despite efforts by some to steal test content or cheat on high-stakes tests? Panelists will look at this issue in three domains:

- Certification/Licensure testing programs
- National testing programs such as the ACT Assessment, SAT, and Graduate Level Testing Programs
- State Educational Assessments

The panelists will cite existing legislation that could be considered by other jurisdictions and suggest areas where there appear to be no relevant laws at this time, but where testing programs would benefit from changes. An example is making cheating by test administrators a crime.

Test-Security Salad: A Test-Security Talk Show (Plenary Session)

Jamie Mulkey, Walt Drane, Victoria Quinn-Stephens, Jennifer Cunningham, & Neal Kingston

A state assessment director, two certification-security managers, and a university assessment director walk into this bar… a salad bar. Join us for Test-Security Salad, a talk show that discusses current issues in test security. Our guest panel will speak about test security from different perspectives. How many of their issues will be the same? Which test-security issues will be different? What is each of their highest test-security priorities? You will learn all this and have an opportunity to ask questions as we present: Test-Security Salad, a test-security talk show.

At the end of this session, the participants will be able to:

- Distinguish between how different types of testing programs implement test security within their organization.
- Gain insight and ideas about how to implement best practices in test security.
- Compare test-security processes and practices discussed against their own program's processes and practices.

Presenters

Steve Addicott
Caveon
steve.addicott@caveon.com

Linda Althouse
American Board of Pediatrics
lalthouse@abpeds.org

Ben Babcock
American Registry of Radiologic Technologists
ben.babcock@arrt.org

Zacch Becker
ProctorU
vtermini+zacch@proctoru.com

Dmitry Belov
Law School Admission Council
dbelov@lsac.org

Kim Brunnert
Elsevier
k.brunnert@elsevier.com

Joe Brutsche
Pearson
joe.brutsche@pearson.com

Brett Chaney
ACT
brett.chaney@act.org

Michael Clifton
ACT
michael.clifton@act.org

Casey Codd
Ascend Learning
casey.codd@ascendlearning.com

Chuck Cooper
IBM
ccooper@us.ibm.com

Nyka Corbin
Graduate Management Admission Council
ncorbin@gmac.com

Michelle Croft
ACT
michelle.croft@act.org

Stephen Cubbellotti
American Board of Internal Medicine
scubbellotti@abim.org

Jennifer Cunningham
University of Phoenix
jennifer.cunningham@phoenix.edu

Kerri Davis
Microsoft Learning Experiences
kerrid@microsoft.com

Walt Drane
Mississippi Department of Education
wdrane@mde.k12.ms.us

Carol Eckerly
University of Wisconsin-Madison
eckerly@wisc.edu

George Eftang
Accenture
george.eftang@accenture.com

Saundra Foderick
Caveon
saundra.foderick@caveon.com

David Foster
Caveon
david.foster@caveon.com

Christy Frederes
Ascend Learning
christy.frederes@ascendlearning.com

John Fremer
Caveon
john.fremer@caveon.com

Chuck Friedman
Professional Examination Services
cfriedman@proexam.org

Bryan Friess
Pearson
bryan.friess@pearson.com

Ruben Garcia
Innovative Exams
ruben@innovativeexams.com

Jennifer Geraets
ACT
jennifer.geraets@act.org

Ardeshir Geranpayeh
Cambridge English
geranpayeh.a@cambridgeenglish.org

Richelle Gruber
Caveon
richelle.gruber@caveon.com

Deborah Harris
ACT
deborah.harris@act.org

Chi-Yu Huang
ACT
chiyu.huang@act.org

Nooree Huh
ACT
nooree.huh@act.org

Benjamin Hunter
Pharmacy Technician Certification Board
bhunter@ptcb.org

Greg Hurtz
PSI
ghurtz@psionline.com

Erika Johnson
Caveon
erika.johnson@caveon.com

Paul Jones
National Association of Boards of Pharmacy
pjones@nabp.net

Neal Kingston
Achievement & Assessment Institute, University of Kansas
nkingston@ku.edu

Jason Kolb
Michigan Department of Education
kolbj1@michigan.gov

Chris Kolhouse
Accenture
chris.kolhouse@accenture.com

Xiang Kong
blcutest@163.com

Sheryl Lazarus
National Center on Educational Outcomes, University of Minnesota
laza0019@umn.edu

Tamara Lewis
Maryland Department of Education
tamara.lewis@maryland.gov

Xin Liu
Ascend Learning
xin.liu@ascendlearning.com

Yang Lu
ACT
yang.lu@act.org

A. Benjamin Mannes
Caveon
ben.mannes@caveon.com

Joy L. Matthews-Lopez
National Association of Boards of Pharmacy
jmatthews-lopez@nabp.net

Dennis Maynes
Caveon
dennis.maynes@caveon.com

Rory McCorkle
PSI
rmccorkle@psionline.com

Tonya Mead
Office of the State Superintendent of Education
tonya.mead@dc.gov

Tara Miller
Ascend Learning
tara.miller@ascendlearning.com

Christine Mills
Ascend Learning
christine.mills@ascendlearning.com

Kathy Moore
Kentucky Department of Education
kathy.moore@education.ky.gov

Keith Morical
Assessment Systems
kmorical@assess.com

Jamie Mulkey
Caveon
jamie.mulkey@caveon.com

Jenifer Mutchie
Caveon
jenifer.mutchie@caveon.com

Ray Nicosia
ETS
rnicosia@ets.org

Victoria Nomdedeu
District of Columbia Public Schools
victoria.nomdedeu@dc.gov

John Olson
Caveon
jmclkolson@yahoo.com

Joslyn Overby
New Mexico Department of Education
joslyn.overby@state.nm.us

Jay Parchure
ACT
jay.parchure@act.org

Marianne Perie
Center for Educational Testing & Evaluation, University of Kansas
mperie@ku.edu

Victoria Quinn-Stephens
Cisco
vquinnst@cisco.com

Brian Reiter
Hawaii Department of Education
brian_reiter/sas/hidoe@notes.k12.hi.us

Amin Saiar
PSI
amin@psionline.com

Karen Schmidt
University of Virginia
kschmidt@virginia.edu

Rachel Watkins Schoenig
ACT
rachel.schoenig@act.org

Marcus Scott
Caveon
marcus.scott@caveon.com

Jennifer Ancona Semko
Baker & McKenzie
jennifer.semko@bakermckenzie.com

William Skorupski
University of Kansas
wps@ku.edu

Nicholas Tardif
Elsevier
n.tardif@elsevier.com

Sarah Thomas
University of Virginia
slt4bg@virginia.edu

Nathan Thompson
Assessment Systems
nthompson@assess.com

Martha Thurlow
National Center on Educational Outcomes, University of Minnesota
thurl001@umn.edu

Jill Van den Heuvel
Alpine Testing Solutions
jill.vandenheuvel@alpinesting.com

Timo von Oertzen
University of Virginia
timo@virginia.edu

Howard Wainer
National Board of Medical Examiners
hwainer@nbme.org

Hongling Wang
ACT
hongling.wang@act.org

Susan Weaver
Caveon
susan.weaver@caveon.com

John Weiner
PSI
john@psionline.com

Marc Weinstein
Dilworth Paxson LLP
mweinstein@dilworthlaw.com

Tara Williams
Caveon
tarawilliams.ce@caveon.com

Leila Williams
Arizona Department of Education
leila.williams@azed.gov

James Wollack
University of Wisconsin-Madison
jwollack@wisc.edu

Karen Wood
Ascend Learning
kwood@atitesting.com

THANK YOU

FOR ATTENDING THE 2015 CONFERENCE ON TEST SECURITY

cete.ku.edu

@CETEmedia