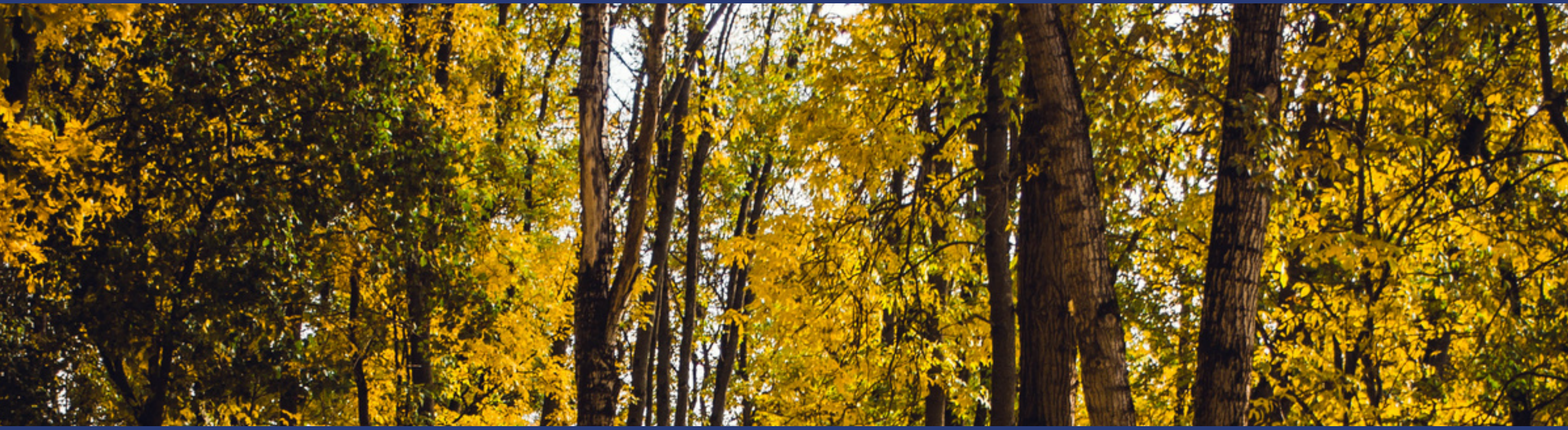
THE 11TH ANNUAL

CONFERENCE ON TEST SECURITY

NOVEMBER 9-10, 2022

VIRTUAL

11:00 A.M. - 4:45 P.M. EDT



A SPECIAL THANKS TO OUR HOST

A SPECIAL THANKS TO OUR CO-HOSTS

ascend LEARNING

caveon™ Test Security

examity® BETTER TEST INTEGRITY.

MEAZURE LEARNING

proctoru yardstick

Pearson VUE

A SPECIAL THANKS TO OUR FRIENDS

ACT®

The AICPA®

Alpine Testing Solutions

CollegeBoard

NCME national council on measurement in education

duolingo english test

ncbe® National Conference of Bar Examiners

National College Testing Association NCTA

PROMETRIC®

psi® Testing Excellence

question mark

WEDNESDAY, NOV 9
11:00 A.M. – 4:45 P.M. EDT

Conference On Test Security 2022

11:00 - 12:00

The Aftermath of Cheating

Camille Thompson, College Board* | *Rachel Schoenig, Cornerstone Strategies* | *Ray Nicosia, ETS* | *Faisel Alam, Law School Admission Council

Panel Presentation

Testing programs face the threat of cheating every day. Incidents can be as small as one individual receiving answers from another test taker or as broad as large, organized groups of individuals intent on stealing test content or sharing answers. Some events can be contained and addressed quickly; others evolve into external scandals that threaten the public's trust in the testing program and the ongoing viability of the organization. During this session, experienced professionals will share case studies involving significant cheating scandals and discuss what comes after cheating. From retiring test items, to cancelling scores, responding to media demands, mending reputational damage, testifying in criminal hearings, and surviving internal shake-ups, participants will discuss the very real impact of cheating and how you and your organization can be better prepared to withstand the aftermath.

WEDNESDAY, NOV 9
11:00 A.M. – 4:45 P.M. EDT

12:15 - 1:15

AI and Data Analytics Approaches to Address Test Security Issues from Remotely Proctored Tests

Jiangang Hao,* *ETS* | *Ikkyu Choi,* *ETS* | *Mo Zhang,* *ETS* | *Paul Deane,* *ETS* | *Chen Li,* *ETS

Coordinated Symposium

COVID19 pandemic accelerated the implementation of remotely proctored high-stake assessments. Complementary to traditional psychometric and statistical analyses, clickstream process data that capture the fine-grained interaction between test takers and items can provide rich diagnostic information for improving test security. In this coordinated session, we gathered three presentations to show how to apply data analytics and AI techniques to clickstream telemetry data and keystroke process data to detect remote desktop access, identify imposters in writing, and identify copy writing behaviors. These three presentations are examples of our comprehensive research agenda on test security based on clickstream data from remotely proctored tests.

1:30 - 2:00

COFFEE CHAT

Join us to discuss themes in the 2019 movie "Setters," shown at the in-person conference. The movie is loosely based on the true events of the Vyapam Scam, an entrance examination scandal that took place in the Indian states of Madhya Pradesh and Uttar Pradesh involving bribery, proxy test taking, item theft, and more.

2:15 - 3:15

Exploration of LOFT and CAT Test Security Analyses and Resilience to Pre-knowledge

Kirk Becker,* *Pearson VUE* | *Huijuan Meng,* *Amazon* | *Jennifer Davis,* *Amazon* | *Mihaiela Gugiu,* *National Registry of Emergency Medical Technicians

Coordinated Symposium

Computer Adaptive Tests and other pool-based exam formats provide increased security relative to fixed test forms, but present challenges for the analysis of collusion due to the limited item overlap between test takers. The relatively large number of items which a pool-based test draws from should also increase the difficulty of memorizing enough content to pass while remaining below minimally competent, and a CAT should partially self-correct when test takers have preknowledge. LOFT provides a consistent item exposure/overlap rate, while CAT overlap and item exposure are conditional on test taker ability and candidate population. This session includes 4 sessions investigating the development and use of test security indicators for a large scale program moving to LOFT, an existing program introducing test security measures for CAT, and simulation research comparing CAT and LOFT security.

WEDNESDAY, NOV 9
11:00 A.M. – 4:45 P.M. EDT

Conference On Test Security 2022

3:30 - 4:45

STANDARD PRESENTATIONS

Privacy and Security: An Unstoppable Force Versus an Immovable Object?

John Kleeman, Questionmark* | *Jamie Armstrong, Questionmark

Which is more important, security or privacy?

If you suspect test fraud, you need to crack down on it effectively. But test takers have privacy rights too—ethically and often legally. This session looks at what happens when security meets privacy. Is privacy an immovable object that trumps security? Or is security an unstoppable force that needs to override privacy? This short session looks at ten questions where security and privacy potentially conflict and offers a one-minute answer on whether security or privacy wins.

Three of the questions to be covered:

1. A test taker asks to have a copy of the answers they gave in the exam. Do you have to provide them?
2. If you catch someone cheating at a test, and as a result, they demand you delete all their personal information under privacy legislation, do you have to delete it?
3. If AI indicates a high probability of cheating, can you stop the exam without a human review?

Join this session presented by two industry practitioners to gain insight into the delicate balance of security and privacy, plus much more.

3:30 - 4:45 CONTINUED...

STANDARD PRESENTATIONS *CONTINUED...*

Investigation of Performances of the Person-Fit Statistics Λ_s and χ_pf for the Lognormal Response Times in Computerized Adaptive Testing

Önder Sünbül, Mersin University* | *Ebru Balta, Agri Ibrahim Cecen University* | *Arzu Uçar, Hakkari University

This study has been aimed to investigate the performance for Λ_s and χ_pf person fit statistics in computerized adaptive testing (CAT). The CAT pool contained 1,500 items. 200 data sets were generated with 100 iterations that response times (RTs) were modeled with the log-normal response time model (LNMRT) under the condition of difficulty level of the compromised items (medium difficulty, difficult) for obtained means of Type I error rate of methods. In order to obtain the means power rate of the methods, 800 data sets were generated with 100 iterations under the conditions of the difficulty level of the compromised items and the ratio of compromised items (20%, 40%, 60%, 80%). Each data set consisted of simulated RT data in 80 items of 10,000 examinees according to LNMRT.The maximum likelihood estimator (MLE) was used to estimate the person speed parameters of the LNMRT for each data set. In the cheating scenario, where item preknowledge examinees were selected from those with low ability level, the rate of fraudulent data was created as 5% of the 10,000 examinees. The RTs of the items of the examinees who had item preknowledge were changed. After this process, we computed the MLEs of the person speed parameters from the (changed) data set. R software was used for data generation and analyses. As a result of the study, it was found that the person fit statistic χ_pf has a high performance in CAT than person fit statistic Λ_s for RT patterns under all conditions.

Boosting Prediction Accuracy for Aberrant Behavior Detection in Standardized Testing by an Enhanced Stacking Machine Learning Algorithm

Todd Zhou, Winston Churchill High School

Test fairness is one important validity issue in large-scale assessment. However, cheating leads to test fairness and security concerns and jeopardizes the validity of test score interpretation. The dramatic increase of online testing has called for more effective and efficient methods to detect cheating. A few studies explored the use of machine learning methods to analyze item responses and response time, but are limited to individual machine learning algorithms or capsulized ensemble learning methods. This study utilized and enhanced the stacking ensemble machine learning algorithm to develop a novel method for cheating detection in standardized testing. Based on the 2-level structure of standard stacking learning method, this study added the outlier-ness information from anomaly detection and the most effective features from the original dataset, into the second level's feature space for the meta-model to learn. The issue of class imbalance in cheating detection was tackled by resampling. The final model's performance was compared with that of the standard stacking learning model and each of the base models, on item responses and response time data. The findings from this study improve the accuracy of aberrant behavior detection in standardized testing and enhance the validity of test score uses.

Simulating the Security Benefits of Randomly Parallel Tests

Andrew Marder, Caveon Test Security

This presentation will address three questions: 1) What is a Randomly Parallel Test (RPT)? 2) How do RPTs improve test security? 3) How much do RPTs improve test security? The bulk of the presentation will focus on the third question, quantifying the security benefits of RPTs. A numerical simulation will provide concrete numerical answers. Like all simulations, the conclusions from this simulation depend upon its assumptions. We aim to make the simulation flexible, accommodating a range of assumptions, and illustrating how our conclusions change with our assumptions. This simulation study compares the effectiveness of an exam delivered using randomly parallel test forms to an exam delivered using a single fixed form. We explore how exam length, item difficulty, item discrimination, size of security breach, and usefulness of pre-knowledge impact reliability and accuracy of rankings for the two exams.

THURSDAY, NOV 10
11:00 A.M. – 4:45 P.M. EDT

Conference On Test Security 2022

11:00 - 12:00

The Big, Bad Wolf in Sheep's Clothing - Protecting Your Program from Remote Access Proxies

Steve Addicott, Caveon Test Security** | **Jon Jensen, Examity** | **Pete Van Dyke, Amazon Web Services** | **Mark Stevens, SAS

Panel Presentation

Since Covid, remote proctoring proved to be a godsend for test programs, allowing many to return to "testing as (nearly) usual" after test centers shut down. However, the rapid adoption of this modality has presented us with an old adversary in a new form—proxies are taking tests by remotely connecting to test takers' computers in manners that are incredibly hard to detect. Indeed, we've arrived at the unfortunate spot where remotely accessing a test taker's computer is now so easy (and profitable), remote proxies utilize social media to not just advertise their services, but even prospect for new business! How can a program protect itself from these wolves in sheeps' clothing?

Two test program leaders and a test security executive will share key lessons learned in confronting these cheaters. They will outline their experiences and foster group discussion with audience members, on these topics:

- Remote-access proxying—how does it work?
- The growing tide of remote access proxies
- Lessons-learned around: 1) Prevention--What can you do to stymie proxies? 2) Deterrence—Risk vs reward, can your program swap the carrot for a stick? 3) Detection and Response--Advances in statistical analysis of test data can identify proxies.

By participating in this discussion, attendees will gain understanding and appreciation of the impacts of a rapidly growing challenge to test score validity. Proven tools to help confront the issue will shared, and best practices for deploying them offered.

12:15 - 1:15

Standards and Best Practices for the Online Observation of Tests

Rachel Schoenig,* *Cornerstone Strategies* | *Mike Murphy,* *ProctorFree* | *Jarret Dyer,* *College of Dupage* | *Stephanie Dille,* *Meazure Learning

Panel Presentation

The National College Testing Association and Association of Test Publishers jointly created standards and best practices for the online observation of tests. The finished document addresses the expectations for test publishers and vendors when delivering assessments using online observation tools, including delivery with or without a proctor. During this session, we will review the process for development of the standards and best practices as well as the intended use of the document. Join members of the technical working group responsible for this document to review key sections dealing with exam security. Together, we will share foundational considerations for testing programs and areas that are continuing to evolve.

MEAZURE
LEARNING
Secure Testing Solutions that are Transforming the Industry.

1:30 - 2:00

COFFEE CHAT

Marc Weinstein, Marc J Weinstein PLLC and Caveon Test Security | Rachel Schoenig, Cornerstone Strategies | Camille Thompson, College Board

Join us to discuss the recent federal court case in Ohio that decided a room scan conducted during a remote proctored online test at a state university amounted to an unreasonable government search in violation of the Fourth Amendment of the United States Constitution.

THURSDAY, NOV 10
11:00 A.M. – 4:45 P.M. EDT

Conference On Test Security 2022

2:15 - 3:15

The Challenges of International Test Security

Camille Thompson, The College Board* | *Nicole Miller, NBME* | *Claire McCauley, Cambridge English Assessments* | *Rachel Schoenig, Cornerstone Strategies
Panel Presentation

Testing in a single jurisdiction has its challenges. Testing across multiple jurisdictions has even more. Varying jurisdictional laws, cultural mores, and technology infrastructure make international exam security more complex and nuanced than ever before. Join testing experts with international exam security experience to explore how to best position your program for success, no matter where it is delivered. We will share real world examples of these challenges and discuss issues ranging from item writing and reviews, translation, and training, to content protection and score validity responses. Together, we will help prepare your organization to better meet the challenges of international testing and test security.

THURSDAY, NOV 10
11:00 A.M. – 4:45 P.M. EDT

3:30 - 4:45

Closing Keynote

James Tunkey, ***I-OnAsia***
Rachel Schoenig, ***Cornerstone Strategies***

For more than two decades, James Tunkey has conducted international investigations on behalf of corporations and testing programs. During that time, he has conducted hundreds of investigations into white collar crimes, test center fraud, and proxy testing rings. This session will explore emerging trends in international fraud, fraud activities impacting testing programs, and key investigative considerations when faced with test fraud. Join us for an engaging and interactive session as we close out COTS 2022!