

SECURITY SIMPLIFIED:

Decoding Data Forensics into Enforceable Evidence

MORE

**Technology in our Measurement
Measurement in our Technology**

Lisa S. O'Leary, Ph.D. – V.P., Product and Service Management
Corina M. Owens, Ph.D. – Manager, Client Solutions

& Senior Psychometricians



Technology-Enabled, People-Powered, Validity-Centered



Test Development & Psychometric Solutions

Customized, technology-enhanced assessment development and consulting services



CertMetrics - Candidate, Credentialing, & Exam Data Management

Integrated credential logic, exam data, and candidate management platform



Security Simplified - Exam & Program Security

Comprehensible prevention, mitigation, detection, and enforcement outcomes



Development Platform & Test Delivery Solutions

Unified ecosystem inclusive of item banking, publication, and delivery options



Prevention



Mitigation



Detection



Enforcement

Systematic security

- defining what preventative actions can be taken to deter anomalous testing behavior,
- defining what data forensic and statistical analyses could be run to address your testing programs unique needs (and how to interpret those findings in a useful way),
- providing guidance on defensible enforcement actions, and
- recommending mitigation strategies to implement at the exam and program level to decrease security issues in the future.

Prevention

- Rapid and robust content development
 - Deep item banks
- Publication strategy and cadence that matches your exam audience and security concerns
- Alignment of policies across your program
- Delivery modality and selection of delivery providers
- SME and system protocols
- Verified credentials



Retake Policy Rules	Failing Outcome Waiting Period	Flags when a candidate violates retake policy for attempt AFTER FAIL within x number of days
	Passing Outcome Waiting Period	Flags when a candidate violates retake policy for retake attempt AFTER PASS within x number of days (RECERTIFICATION POLICY)
	Pass in Perpetuity	Flags when a candidate violates retake policy for any attempt AFTER PASS (EXAM STATUS FOR LIFE)
	Total Attempts in Time Limit	Flags when a candidate violates retake policy for total number of attempts exceeds x within z timeframe

Mitigation

- Active monitoring of item, form, and candidate performance trends throughout the life cycle of exams and programs
 - Same exam across time
 - Review longitudinal data to determine exam specific next steps
 - Same program across exams
 - Review aggregate data to discuss programmatic patterns



Mitigation: Web Crawling

- Aides in the monitoring of stolen exam content that is available on the web
 - Availability and amount of exam content available on the web is clear and obvious evidence that content has been exposed
- May be used in conjunction with forensic analyses or used independently



Mitigation

- Warning lights activated, now what?
 - Conduct targeted data forensics based on security concern
 - Revisit prevention strategies to take exam specific action
 - Content development
 - Form publication
 - Exam delivery
 - Adjust and communicate
 - Candidate policies
 - Delivery provider agreements
 - Test design details





Data Forensics: CertMetrics Security Scripts – Piracy and Access to Content

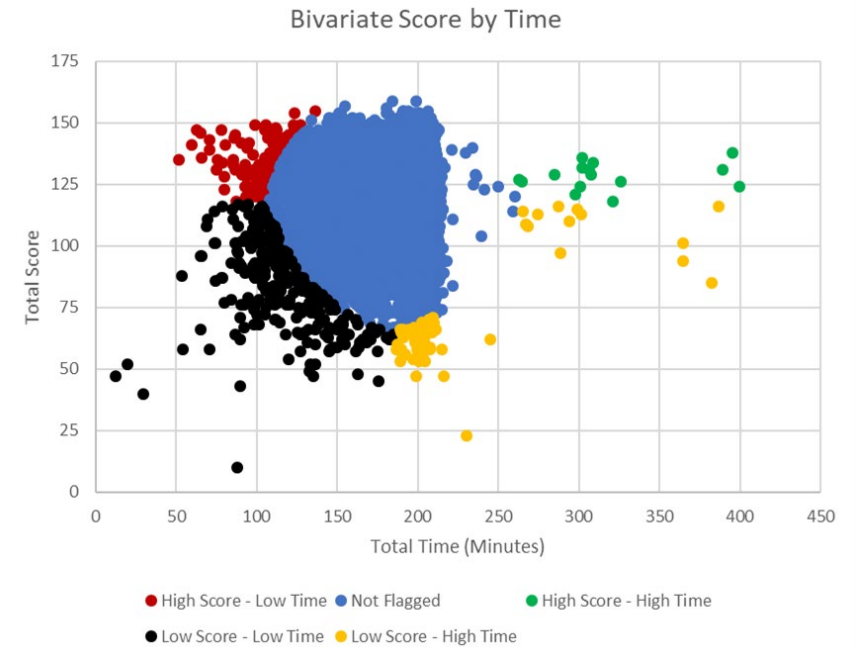
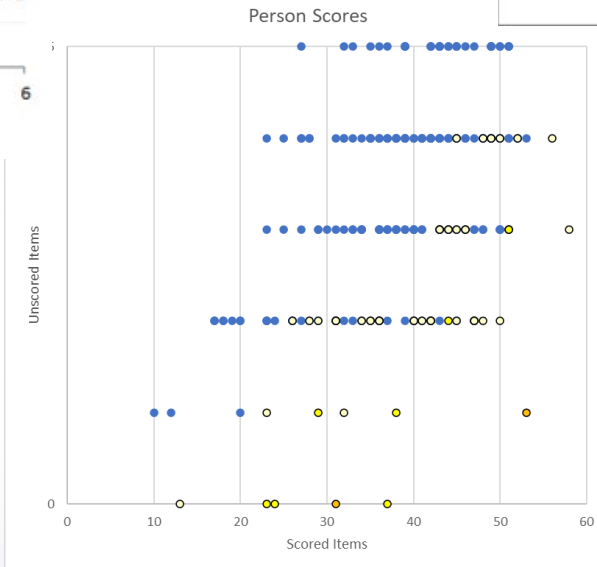
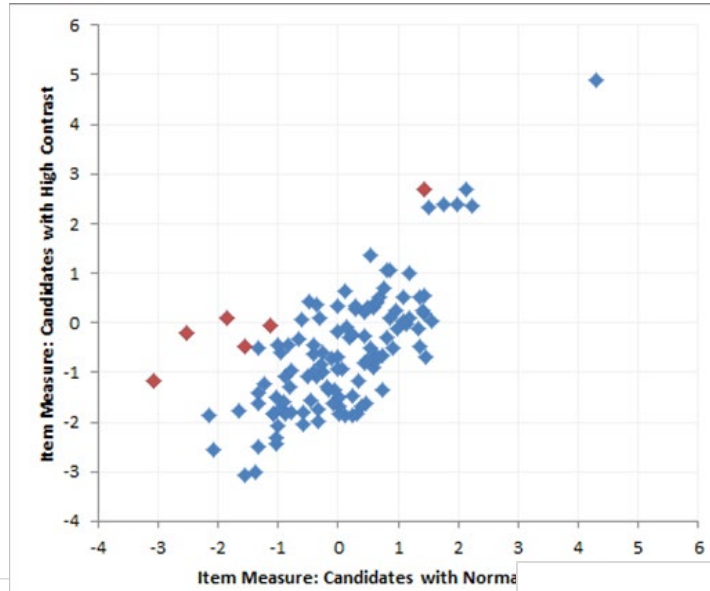
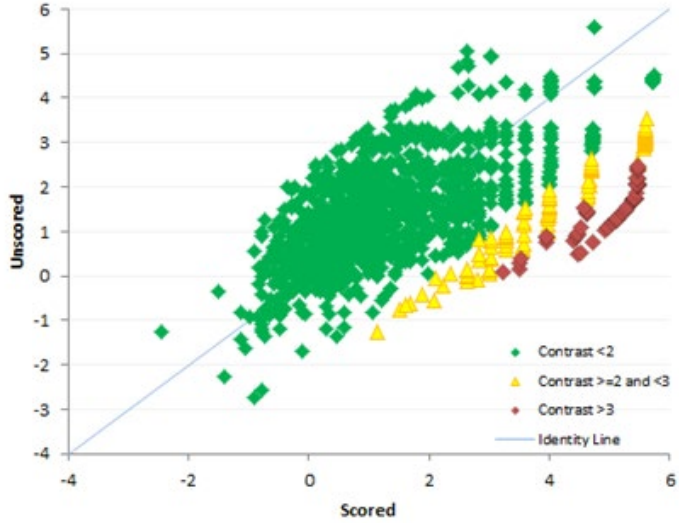
Exam Time Rules	Item Skipped or Answered Too Quickly	x% of items skipped or answered in less than y seconds
	Exam Passed Too Quickly	Scoring more than x% in less than y minutes
	Exam Taken Too Quickly	Less than y minutes
Exam Form Rules	Exam retake score increase too high	x% score increase in y days
Item Performance Comparison Rules	Item Subset Differential	x% or more for Subset 1 & y% or less for Subset 2

Data Forensics – Piracy and Access to Content

- Identification of Candidates' Pre-Knowledge
 - Differential Person Functioning (DPF)
 - Bivariate Score by Time (BST)
- Identification of Exposed Content
 - Differential Item Functioning (DIF)
 - Drift
 - Unscored-Only Analysis



Candidate Measures on Scored and Unscored Items



Data Forensics – Candidate Conduct

- Collusion among test takers
 - Response Similarity Index (RSI) Analyses
 - Score Similarity Index (SSI) Analyses
 - Cluster Analyses
- Proxy Test Taking
 - Flags from delivery provider
 - Behavior analyses
 - Proctoring information
 - Collusion analyses



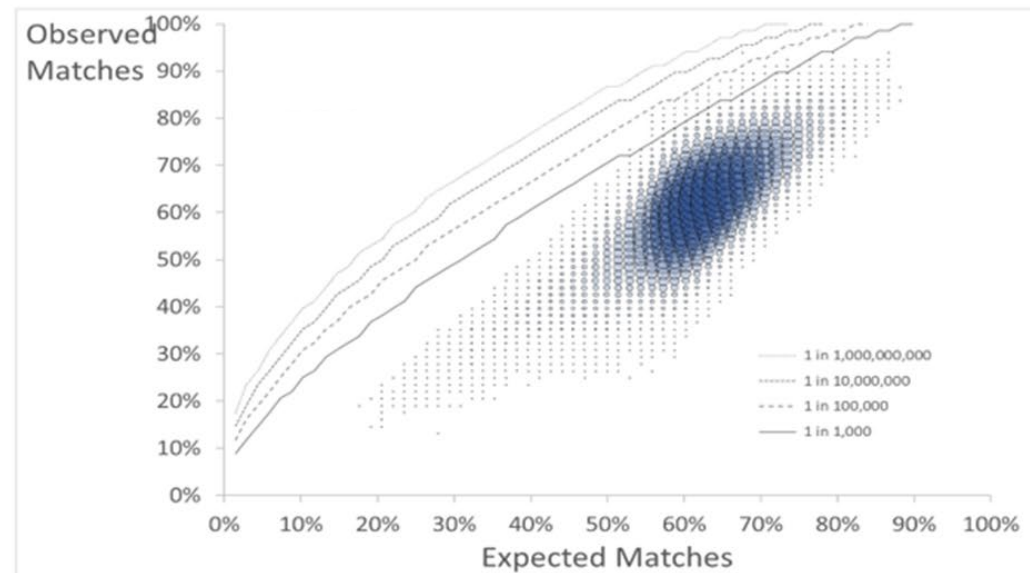
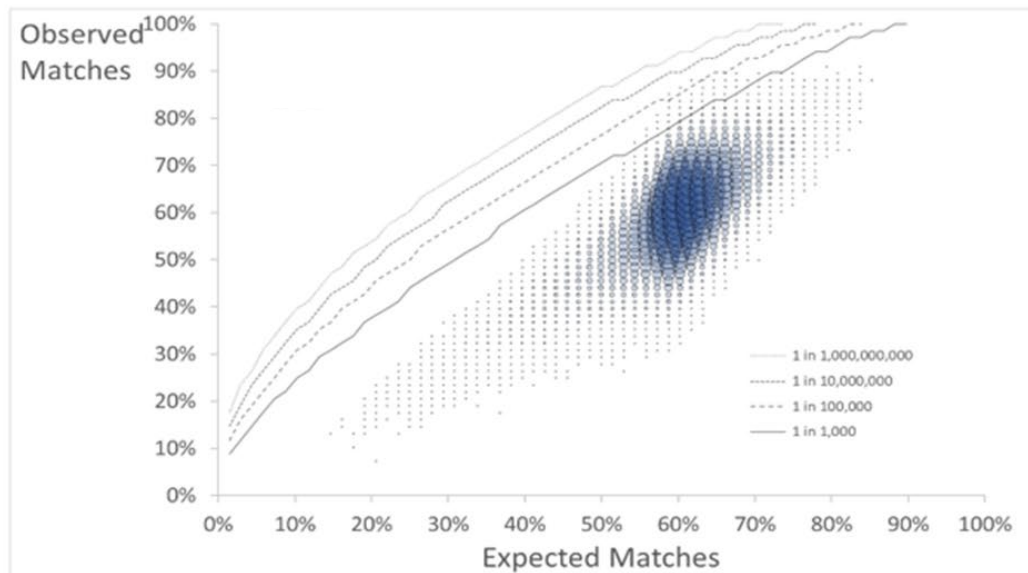
Data Forensics – Candidate Conduct



Item Performance Comparison Rules	Approximation Score Similarity Index (SSI)	Z-score threshold + # of matches
	Approximation Response Similarity Index (RSI)	Z-score threshold + # of matches

Similarity Indices with Clustering

Form	Total No. of Candidate Records	Total No. of Pairs Analyzed	No. of Flagged Pairs	Percent Flagged Candidate Pairs	No. of Unique Candidate Records Flagged	Percent Flagged Candidate Records	No. of Clusters
A	4,883	11,919,403	536	0.004%	281	5.75%	97
B	4,828	11,652,378	323	0.003%	247	5.12%	94





Criteria for Describing Likelihood of Involvement in Other Suspicious Activity

Extreme

- Extreme RSI/SSI p-values <1.0E-07, 3+ BST Flags (HS/LT)
- Extreme RSI/SSI p-values,

Very High

- RSI and BST (HS/LT) flag
- Extreme SSI p-value
- Extreme SSI p-value and RSI flag

High

- RSI and BST (HS/LT) flag
- 3+ <20 Sec and BST Flags, 5+ SSI Flags
- 10+ SSI flags, min SSI p-value <1.0E-05
- 10+ SSI flags, RSI Flag
- 10+ SSI flags, BST or <20 Sec flag
- 5+ SSI flags, min SSI p-value <1.0E-05, RSI Flag
- 5+ SSI flags, min SSI p-value <1.0E-05, and BST/<20 sec flag
- Multiple RSI Flags

Moderate

- 3+ BST Flags (HS/LT)
- 3+ <20 Sec and BST Flags
- 10+ SSI flags
- 5+ SSI flags, min SSI p-value <1.0E-05

Level of Suspicion	Date	Testing Center	
(0) Moderate	8/11/2021	xz0x	
(0) Moderate	5/1/2021	5	
(0) Moderate	10/24/2021	z6	
(0) Moderate	11/8/2021	5zx6	
(0) Moderate	10/16/2021	6x	
(1) High	3/3/2021	9y0z	
(0) Moderate	3/8/2021	9y0z	
(2) Very High	5/4/2021	x6x8	
(0) Moderate		Extreme RSI/SSI p-values	(3) Extreme
(0) Moderate		5+ SSI flags, min SSI p-value <1.0E-05, RSI Flag	(1) High
(0) Moderate		10+ SSI flags, min SSI p-value <1.0E-05	(1) High
(0) Moderate		10+ SSI flags, min SSI p-value <1.0E-05	(1) High
(0) Moderate		10+ SSI flags	(0) Moderate
(0) Moderate		10+ SSI flags	(0) Moderate
(0) Moderate		10+ SSI flags	(0) Moderate
(0) Moderate	6/20/2021	9y06	

	Written	Decimal Notation	Scientific Notation	Excel Scientific Notation
1 in 10	one in ten	0.1	10 ⁻¹	1.0E-01
1 in 100	one in one hundred	0.01	10 ⁻²	1.0E-02
1 in 1,000	one in one thousand	0.001	10 ⁻³	1.0E-03
1 in 10,000	one in ten thousand	0.0001	10 ⁻⁴	1.0E-04
1 in 100,000	one in one hundred thousand	0.00001	10 ⁻⁵	1.0E-05
1 in 1,000,000	one in one million	0.000001	10 ⁻⁶	1.0E-06
1 in 10,000,000	one in ten million	0.0000001	10 ⁻⁷	1.0E-07
1 in 100,000,000	one in one hundred million	0.00000001	10 ⁻⁸	1.0E-08
1 in 1,000,000,000	one in one billion	0.000000001	10 ⁻⁹	1.0E-09

Enforcement

- Consult with legal
- Establish written security policy
- Require candidate agreements
- Establish candidate appeals process
- Conduct comprehensive data forensics
- Triangulate multiple sources evidence of anomalous behavior prior to taking action





Enforcement with CertMetrics

Candidate Access Rules	Watch List Test Taker	Flags at the time of import when a candidate that is on an established Watch List takes an exam (administration date)
	Banned Candidate Test Taker	Flags at the time of import when a candidate that is on an established Banned List takes an exam (administration date)
	Prohibited Country Test Taker	Flags at the time of import when a candidate, with a home address in a client-specified prohibited country (by ISO verified country)

Enforcement Actions

- Warning email
- Require review prior to exam/certification results being available
- Restriction on future registration
- Exam status change
- Credential status change
- Add to watch list
- Add to banned list



- Document security prevention, mitigation, detection, and enforcement policies
- Consult with your legal team/advisors on planned security policies and prior to taking any actions against candidates, test centers based on data forensic results
- Be transparent at a high-level with candidates regarding security policies
- Align candidate agreements with documented policies and processes
- Provide candidates the opportunity and means to challenge resulting actions
- Triangulate multiple sources of evidence of anomalous behavior
- Treat all candidates equally (e.g., do not target data forensic techniques at a particular subset of candidates)
- Continually evaluate of exam and program to assess appropriate actions



SECURITY

Simplified

lisa.oleary@alpinetesting.com