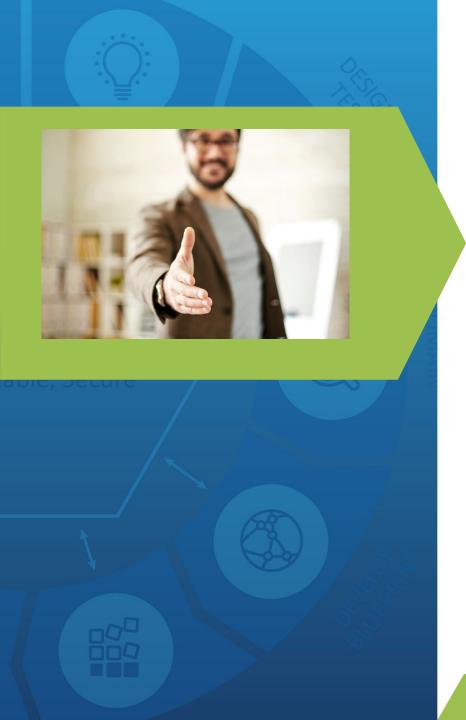


Summer Security Webinar: Part 2

Supporting certification programs through the power of CertMetrics

Presenters:

Brett Foley, Ph.D., Psychometric Services and Industry Practice Lead Tara Miller, Technical Solutions Designer





- Introductions
 - Brett Foley, Ph.D. Psychometric Service & Industry Practice Lead
 - Tara Miller Technical Solution Designer

ALPINE'S CORE COMPETENCIES



TEST DEVELOPMENT

PSYCHOMETRIC CONSULTING



CANDIDATE

CREDENTIALING

EXAM DATA
MANAGEMENT

Certification Programs Come in a Variety of Sizes







Certification – Many systems and vendors Vendor Management Complex item types roduct Enhancements **Company Targets**









Reason IRT is awesome #7: People and Items on Same Scale!



Brett

• Ability = Low (-1.5)



Tara

• Ability = High (2.4)



Question

• Difficulty = Medium (0.5)

Which of the following is an example of a nuclear fission reaction?

$$_{A,4}^{8}$$
Be + $_{2}^{4}$ He $\longrightarrow _{6}^{12}$ C

$$_{\text{B.}} \stackrel{234}{90}\text{Th} \longrightarrow {}^{234}_{91}\text{Pa} + {}^{0}_{-1}e$$

$$_{C_{1}}^{27}\text{Al} + {}_{2}^{4}\text{He} \longrightarrow {}_{15}^{30}\text{P} + {}_{0}^{1}n$$

$$D_{0.94}^{239}$$
Pu + ${}^{1}_{0}n \longrightarrow {}^{137}_{52}$ Te + ${}^{100}_{42}$ Mo + $3{}^{1}_{0}n$

Reason IRT is awesome #8: Probabilistic estimates!



Which of the following is an example of a nuclear fission reaction?

$$A._{4}^{8}Be + _{2}^{4}He \longrightarrow _{6}^{12}C$$

$$B._{90}^{234}Th \longrightarrow _{91}^{234}Pa + _{-1}^{0}e$$

$$C._{13}^{13}Al + _{2}^{4}He \longrightarrow _{15}^{30}P + _{0}^{1}n$$

$$D._{94}^{239}Pu + _{0}^{1}n \longrightarrow _{52}^{137}Te + _{42}^{100}Mo + 3_{0}^{1}n$$



Brett

- Ability = Low (-1.5)
- Prob. of correct answer = 12%
- Prob. of incorrect answer = 88%



Tara

- Ability = High (2.4)
- Probability of correct answer = 87%
- Prob. of incorrect answer = 13%

Question

- Difficulty = Medium (0.5)
- Probability Brett & Tara BOTH get the question RIGHT:
 - 12% x 87% = 10%
- Probability Brett & Tara Both get the question WRONG:
 - 88% x 13% = 11%
- Probability Brett & Tara Both get the SAME OUTCOME:
 - **10%** + 11% = 21%

Reason IRT is awesome #9: Security Analysis



Exam Item	Probability of same outcome	
1	22%	
2	46%	
3	29%	
4	41%	
5	20%	
6	1%	
7	40%	
8	26%	
9	31%	
10	44%	
TOTAL	300% or 3	

- For this 10 item quiz:
 - We expect Brett and Tara to have about 3 outcomes in common
 - What if we observe Brett and Tara's Results and they have 8 outcomes in common?
 - There's a 0.2% chance Brett and Tara have 8 or more common outcomes given their ability and the difficulty of the test (about 1 in 600)













Enforcement

Two Alpine Security Tools: SSI & RSI



- Score Similarity Index (SSI)
 - Estimates the probability that the observed similarity in the **scored** responses between any pair of examinees completing the same test form is greater than the expected similarity
- Response Similarity Index (RSI)
 - Estimates the probability that the observed similarity in the actual responses between any pair of examinees completing the same test form is greater than the expected similarity

Big Idea:

Identify pairs of examinees with unusual numbers of matching scores/responses.

SSI & RSI: What are they good for?



Can Identify...

- Pairs of colluders
 - "Sneaky peekers"
- Groups of colluders/exposed content
 - Multiple examinees accessing same brain dump website during preparation.
 - Group of examinees all using the same proxy test takers.

SSI & RSI: Benefits and Limitations



- Score Similarity Index (SSI)
 - Benefits
 - Easier calculation
 - Work for more item types
 - Limitations
 - Potentially decreased certainty

Big Idea:

Both are awesome tools. Best choice will vary based on design of exams/program.

- Response Similarity Index (RSI)
 - Benefits
 - Potential for increased certainty
 - Limitations
 - Works with fewer item types (4/5 option MC)
 - Model assumptions/fit
 - Software/sample size requirements

Limitations of SSI/RSI as stand-alone tools Running occasionally with other psychometric analyses



- Stand-alone SSI/RSI analysis done regularly...
 - Excellent for monitoring security concerns
 - Large number of additional supporting analyses available
- Delayed identification
 - Credentials may have been awarded
 - Harm done
- Requires manual action across systems
 - Result review/decision making
 - Application
 - Delivery
 - Reporting



Solution: CM Advanced Security Scripts



- Approximations of SSI/RSI are built in to CertMetrics
- Not reliant on IRT person/item parameters
 - Closely approximates IRT-based estimates
- Benefits
 - Taking action
 - Analyzed on import
 - Potential to find before credentialing
 - Customizable look-back window
 - Customizable flagging (can be linked to other security indicators)
 - Automation of Actions
 - Complex programs
 - Standardize across multiple credentials
 - Or customize actions for each







Custom Criteria Applied

- Foundational Flags
 - Exam retake policy violation
 - Exam retake after passing
 - Exam retake score increase
 - Items answered too quickly
 - Test taken too quickly
 - Test passed too quickly
 - Banned candidate
 - Embargoed country test taker

- Advanced Security Analyses
 - Approximation Score Similarity Index (SSI)
 - Approximation Response Similarity Index (RSI)
 - Item Comparison Subset Differential









Work Flow

Consultation & Initial Set-up

Data Imports
During
Administrations

Analyses & Flags Applied on Import

Custom Criteria Applied Action
Taken
Automatically

Why are these important? Can't we just use the best settings?

Benefits of Client Consultation



Security approach should match client goals

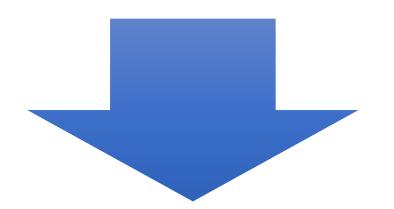
Example goals:

• "I don't know if we have any security problems. Can our exam data help us understand if we do?"

 "We must avoid giving credentials to examinees who cheat. Catching candidates with untrustworthy results and taking action is our priority."

Benefits of Client Consultation Balancing false positives/negatives





More conservative flagging...

- Catches fewer cheaters
- Can increase certainty
- Increases chance for false negatives

More aggressive flagging...

- Catches more cheaters
- Can decrease certainty
- Increases chance for false positives



Client Consultation Example

AlpineTesting Solutions

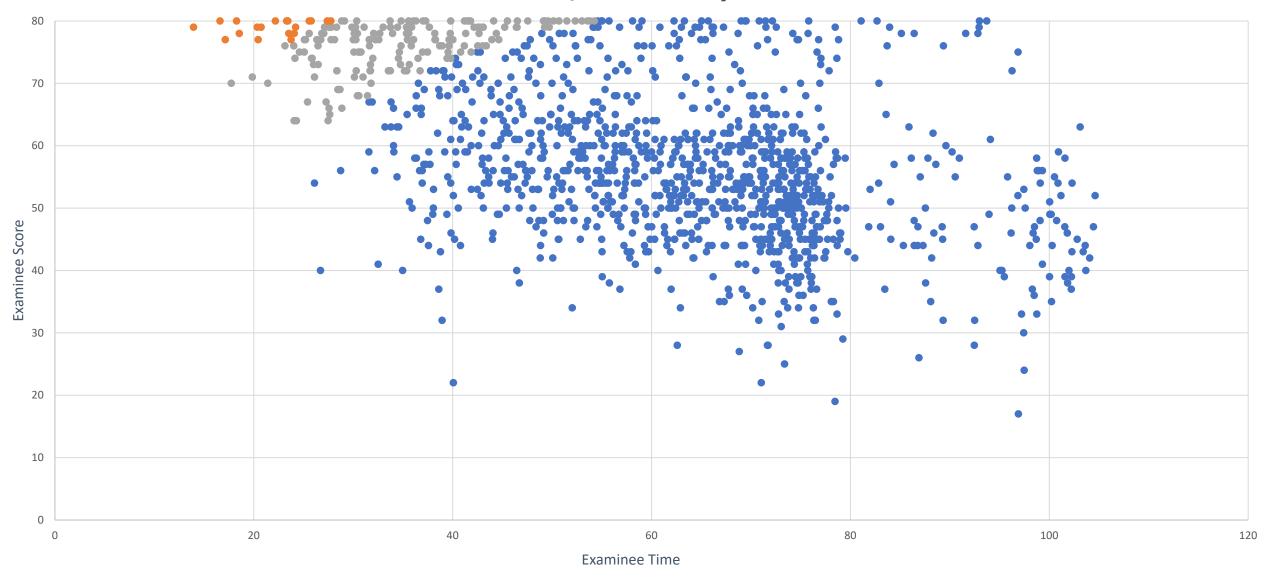
Addressing the probabilistic nature of flagging criteria

- SSI/RSI are probabilistic
- Threshold for "unusual/suspect" must be chosen
 - 1 in 1,000?
 - 1 in 10,000?
 - 1 in 1,000,000,000?

- Number of candidate pairs grows geometrically with the number of exam administrations:
- 100 exams = 4,950 pairs
- 500 exams = 124,750 pairs
- 2000 exams = 1,999,000 pairs

If we use a flagging criteria of 1 in 10,000 and have 2000 exams, we can expect about **200 candidate pairs** to be flagged, just by chance.

Examinee Score/Time Analysis



Client Consultation Example

Flagging only for a trial period



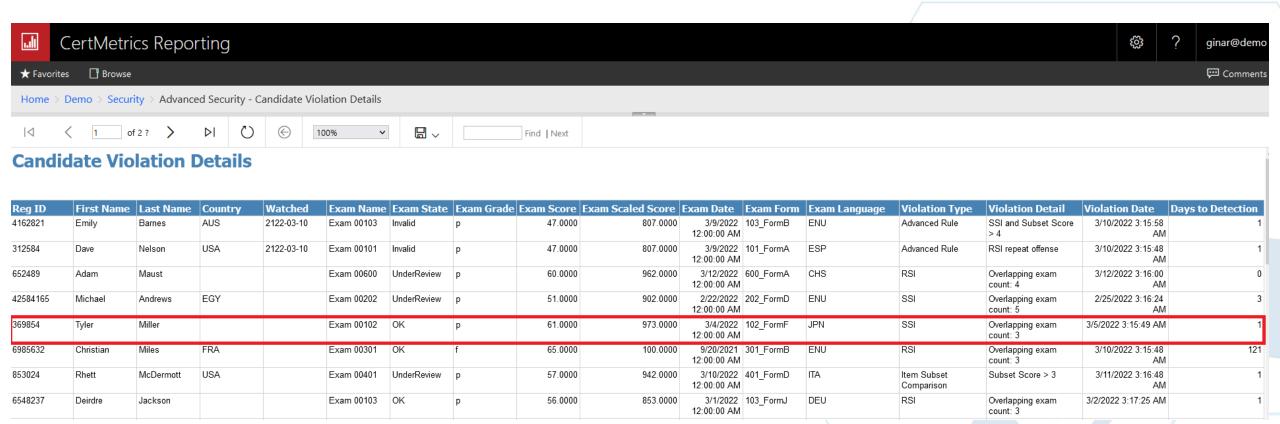
- Can be useful to try out a flagging strategy
 - See how it works (e.g., flagging rate)
 - Make adjustments
- Once flagging criteria have been reviewed, then define automated actions.

Automated Actions

 One of the most powerful features of CM combined with the Advanced Security Module.

CertMetrics Flagged Result - Reporting





.





CertMetrics 2.0
Functionality to
Support your Exam
Security Program

Candidate Account Identity Control
Psychometrically applied flagging
Custom logic and eligibility
Review queues

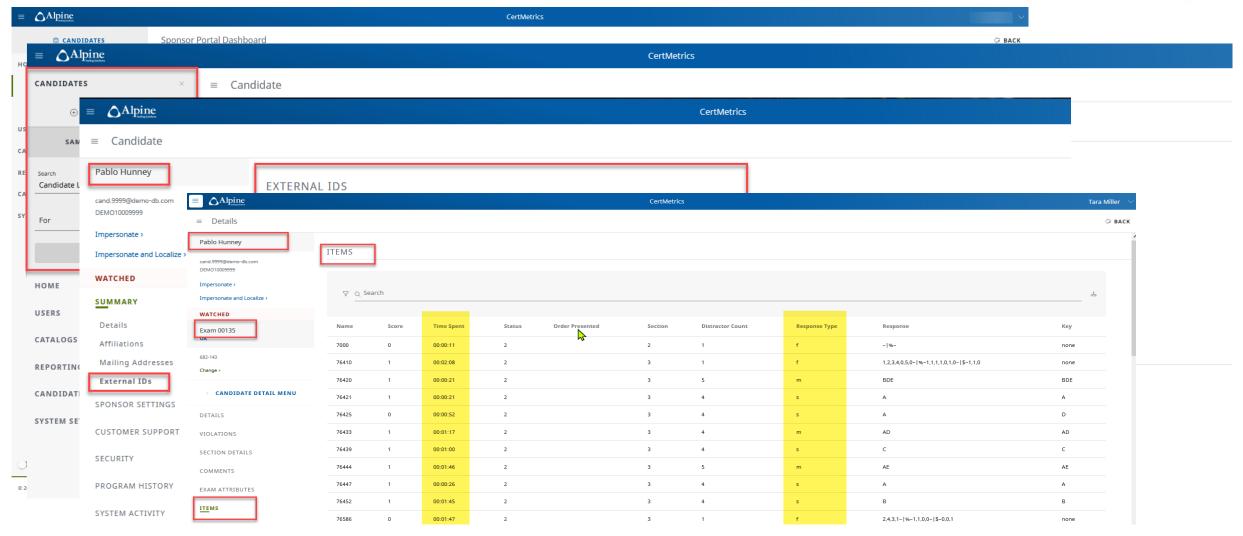
Security review upon import

Automated candidate communications



Inside CertMetrics 2.0

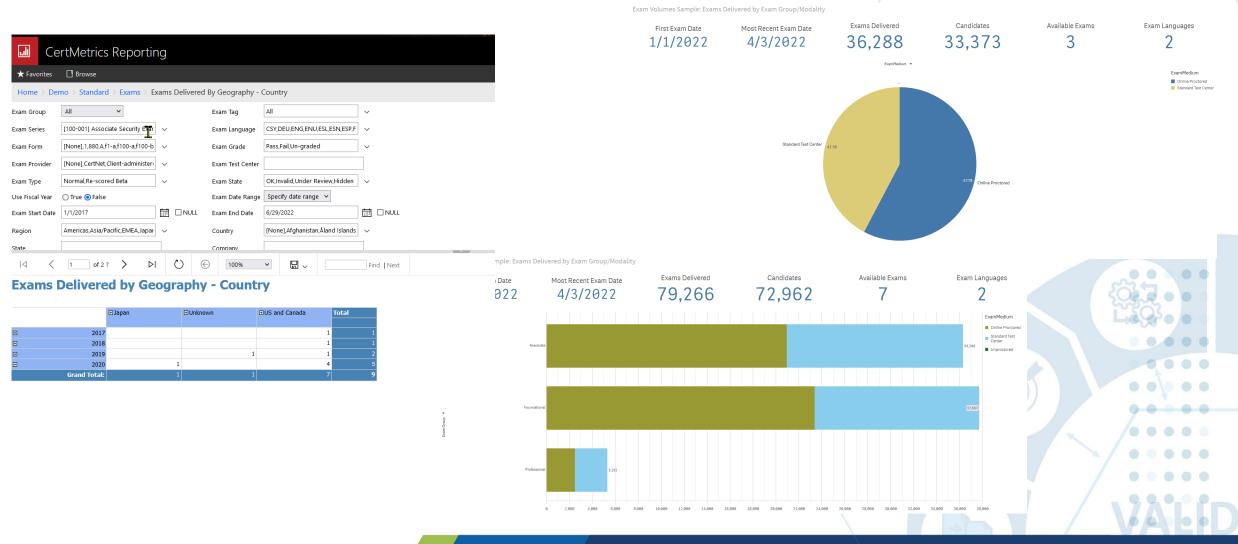




CertMetrics Reporting



SSRS Reports



Qlik Reports





Work Flow

Consultation &
Initial Set-up

Data Imports
During
Administrations

Analyses Applied on Import

Custom Criteria Applied Action Taken Automatically



Thank you for joining us! Questions??

VALIDITY
Fair, Reliable, Secure