Sharing and Comparing:
Best Practices from Education & Credentialing Contexts

Tony Alpert, SBAC
Wayne Camara, ACT
Gregory J. Cizek, UNC
Liberty Munson, Microsoft
Jamie Mulkey (moderator), Caveon
Why share?

- Best practices that should be considered
- Good to get outside of your domain
- Helps to understand different needs for different environments
- We all have the same goal in mind
Today’s session

- Format
- Business Drivers
- Achievement Levels
- Alignment
- Test Length
- Test Security
Business Drivers

Liberty J. Munson
Learning for the Modern Era
Enabling the next generation of innovators, entrepreneurs and developers

<table>
<thead>
<tr>
<th>1</th>
<th>Build Content with Strategic Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>edX MIT Massachusetts Institute of Technology</td>
<td></td>
</tr>
<tr>
<td>Microsoft Partner</td>
<td></td>
</tr>
<tr>
<td>THE LINUX FOUNDATION</td>
<td></td>
</tr>
<tr>
<td>OpenEdx</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2</th>
<th>Extend our Platform through Open Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empower Students &amp; Educators</td>
<td></td>
</tr>
<tr>
<td>Activate Cloud &amp; Productivity Partners</td>
<td></td>
</tr>
<tr>
<td>Enable Customers &amp; Partners at Scale</td>
<td></td>
</tr>
<tr>
<td>Drive Usage, Deployment &amp; Consumption</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3</th>
<th>Go where the Learners are</th>
</tr>
</thead>
<tbody>
<tr>
<td>edX Learning</td>
<td></td>
</tr>
<tr>
<td>Microsoft Partner Learning</td>
<td></td>
</tr>
<tr>
<td>Harvard/MSFT CS50 AP</td>
<td></td>
</tr>
<tr>
<td>C# for Absolute Beginners</td>
<td></td>
</tr>
<tr>
<td>Hour of Code</td>
<td></td>
</tr>
<tr>
<td>Minecraft</td>
<td></td>
</tr>
<tr>
<td>Code York</td>
<td></td>
</tr>
<tr>
<td>LinkedIn</td>
<td></td>
</tr>
<tr>
<td>Microsoft Imagine</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4</th>
<th>Priorities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microsoft Azure Office 365 Windows</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5</th>
<th>Goals and Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total reach</td>
<td></td>
</tr>
<tr>
<td>New/Competitive Developers reached</td>
<td></td>
</tr>
<tr>
<td>Skilled technologists building on Microsoft platforms and services</td>
<td></td>
</tr>
<tr>
<td>Preference for Microsoft platforms &amp; services</td>
<td></td>
</tr>
</tbody>
</table>
An Overview of MS Certification Exams

- Computer administered
- Test center deliveries or online proctoring
- Global distribution
- Ongoing delivery
- Variety of item types (e.g., multiple choice, drag and drop, active screen, hot area, case studies, labs, text entry, code analysis, etc.)
Microsoft’s Certification Structure

The five new expert certifications are:

- MCSE: Cloud Platform and Infrastructure (Windows Server and Microsoft Azure)
- MCSE: Mobility – (Windows Client and Enterprise Mobility Suite)
- MCSE: Data Management and Analysis (on-premises and cloud-based Microsoft data products and services)
- MCSE: Productivity (Office 365, SharePoint, Exchange, and Skype for Business)
- MCSD: App Builder (Web and Mobile app development)

To earn each of these credentials:

- Earn Microsoft Certified Solutions Associate (MCSA) certification
- Pass a single additional exam from a list of electives

Certifications will include achievement date that signifies candidates investment in staying up to date on the technology

- Every year, re-earn the certification by passing an additional exam from the list of electives or retaking an exam on a rapidly changing technology
- Purpose for “recertification”?
  - Breadth
  - Continual competence
  - Engagement
Microsoft Learning Experiences

Drive adoption of Microsoft technologies through training and certification

MOBILITY
- MOBILITY
- MOBILITY
- MCSE Mobility Earned: 2016
  - MOBILITY
  - MOBILITY
  - MCSE Windows 10 Elective
  - MOBILITY
  - MOBILITY
  - MCSE Windows Server 2012
  - MOBILITY
  - MOBILITY
  - MCSE Windows Server 2016
  - MOBILITY
  - MOBILITY
  - MCSE Linux on Azure
  - MOBILITY
  - MOBILITY
  - MCSE Cloud Platform

CLOUD PLATFORM & INFRASTRUCTURE
- CLOUD PLATFORM & INFRASTRUCTURE
- CLOUD PLATFORM & INFRASTRUCTURE
- MCSE Cloud Platform & Infrastructure Earned: 2016
  - CLOUD PLATFORM & INFRASTRUCTURE
  - CLOUD PLATFORM & INFRASTRUCTURE
  - MCSE Windows 10 Elective
  - CLOUD PLATFORM & INFRASTRUCTURE
  - CLOUD PLATFORM & INFRASTRUCTURE
  - MCSE Windows Server 2012
  - CLOUD PLATFORM & INFRASTRUCTURE
  - CLOUD PLATFORM & INFRASTRUCTURE
  - MCSE Windows Server 2016
  - CLOUD PLATFORM & INFRASTRUCTURE
  - CLOUD PLATFORM & INFRASTRUCTURE
  - MCSE Linux on Azure
  - CLOUD PLATFORM & INFRASTRUCTURE
  - CLOUD PLATFORM & INFRASTRUCTURE
  - MCSE Cloud Platform

PRODUCTIVITY
- PRODUCTIVITY
- PRODUCTIVITY
- MCSE Productivity Earned: 2016
  - PRODUCTIVITY
  - PRODUCTIVITY
  - MCSE Office 365 Elective
  - PRODUCTIVITY
  - PRODUCTIVITY
  - MCSE SQL Server 2012/2014
  - PRODUCTIVITY
  - PRODUCTIVITY
  - MCSE SQL 2016 Database Admin
  - PRODUCTIVITY
  - PRODUCTIVITY
  - MCSE SQL 2016 BI Development

DATA MANAGEMENT & ANALYTICS
- DATA MANAGEMENT & ANALYTICS
- DATA MANAGEMENT & ANALYTICS
- MCSE Data Mgmt & Analytics Earned: 2016
  - DATA MANAGEMENT & ANALYTICS
  - DATA MANAGEMENT & ANALYTICS
  - MCSE SQL Server 2012/2014
  - DATA MANAGEMENT & ANALYTICS
  - DATA MANAGEMENT & ANALYTICS
  - MCSE SQL 2016 Database Admin
  - DATA MANAGEMENT & ANALYTICS
  - DATA MANAGEMENT & ANALYTICS
  - MCSE SQL 2016 BI Development

APP BUILDER
- APP BUILDER
- APP BUILDER
- MCSD App Builder Earned: 2016
  - APP BUILDER
  - APP BUILDER
  - MCSE Web Applications Elective
  - APP BUILDER
  - APP BUILDER
  - MCSE Universal Windows Platform
Choose two from:
- 532: Developing Microsoft Azure Solutions
- 533: Managing Microsoft Azure Infrastructure Solutions
- 534: Architecting Microsoft Azure Solutions
- 473: Designing and Implementing Cloud Data Platform Solutions
- 475: Designing and Implementing Big Data Analytics Solutions

**Elective Exam Pool (January 2017)**
- 532: Developing Microsoft Azure Solutions
- 533: Managing Microsoft Azure Infrastructure Solutions
- 534: Architecting Microsoft Azure Solutions
- 473: Designing and Implementing Cloud Data Platform Solutions
- 475: Designing and Implementing Big Data Analytics Solutions
- 413: Designing and Implementing a Server Infrastructure
- 414: Implementing an Advanced Server Infrastructure
- 246: Monitoring and Operating a Private Cloud
- 247: Configuring and Deploying a Private Cloud
Certification & Licensure

- Purpose of certification/license
  - Protect public
  - Drive adoption
  - High, medium, low stakes
- Business goals
  - Reach, revenue, strategic
  - Drive engagement
- Development and sustainment costs
  - Balancing psychometric needs with business reality
- Global vs local distribution
  - Exam delivery providers (EDPs)
- Exam availability
  - Any time, any place delivery
  - Online proctoring options
- Practice tests, preparation materials
  - Align to training?
  - Accreditation
- Security
  - Braindump sites, proxy testers, cheating
  - Proctoring
  - Identity verification
- Scoring
  - Immediate vs delayed
- Role of marketing
  - Value of certification to candidate and employer
  - Competitive space
Business Drivers
Wayne Camara
Business Drivers - Educational Assessments

**B2C (SAT, ACT, GRE, etc.)**
- Access - ability to take test at convenience
- Demand by external agency (college, scholarship service, NCAA)
- Security, turnaround time, repeat testing
- Practice test - How to improve total score
- Standardization, fairness
- Shorter testing time

**B2B - Large scale state testing**
- Cost to state/district
- Alignment to standards
- Instructional sensitivity (subscores & diagnostics)
- Growth over time (across grades)
- Turnaround time
- Administrative flexibility - at school level, state accommodations
- Opt-out
- Shorter testing time.
Business - Drivers in Digital vs paper

B2C (SAT, ACT, GRE, etc.)
- High security - organized efforts to expose content
- High priority on standardization
- Global vs state specific
- Large item banks and rotating forms
- Access to test - local administration, local school
- Saturday administration (Summer date?)
- Paper and Digital versions existing side by side for a long time with strict score comparability
- Device and mode should be no factor in score
- Fee waivers - 15-18% drive costs
- Demand for comparability prevents some innovation - modularity, TEI, innovation, shorter adaptive testing

B2B - Large scale state testing
- Low cost
- High priority for flexibility support (long windows, scheduling, retesting)
- Multiple devices (Chromebook, tablets, laptops) and comparability
- In school testing
- Innovative item types - performance tasks
- Released items and pools for practice
- Rapid score turnaround vs testing late in school year
- Reduced testing time
- Performance tasks, TEI, highly desirable
## Business Drivers - Educational vs Credentialing

<table>
<thead>
<tr>
<th>Category</th>
<th>Educational</th>
<th>Credentialing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Business model</strong></td>
<td>more customer sensitivity</td>
<td>less competition and are often monopolies</td>
</tr>
<tr>
<td><strong>Cost</strong></td>
<td>high sensitivity</td>
<td>less sensitivity</td>
</tr>
<tr>
<td><strong>Test site</strong></td>
<td>prioritizes local administration</td>
<td>Use test center model, requiring travel greater travel/less convenience is a given</td>
</tr>
<tr>
<td><strong>Purpose/Use</strong></td>
<td>Multiple</td>
<td>Single</td>
</tr>
<tr>
<td><strong>Reporting</strong></td>
<td>Student, class, teacher, school, state</td>
<td>test taker</td>
</tr>
<tr>
<td><strong>Security</strong></td>
<td>Site, score, growth focus</td>
<td>Individual, item bank, item pool</td>
</tr>
<tr>
<td><strong>Digital issues</strong></td>
<td>administrative flexibility, connectivity, device familiarity, innovative item types</td>
<td>standardization, single device, items which best represent construct (vs innovation for the sake of innovation)</td>
</tr>
</tbody>
</table>
What Credentialing Can Learn for Educational Business Drivers

- Customer focus - more score information to help test taker improve, identify strengths and weaknesses (GMAT or LSAC vs GRE)
- Security focused on score integrity vs IP
- Cost containment - value add of enhancements vs cost to test taker
- Fee waivers - when credential is a barrier to employment or advancement?
- Transparency - data (subgroup, volumes, validity, reliability)
- Investment in assessment - new forms, items
- Test prep and practice materials - free and accessible
Achievement Levels

Tony Alpert
Achievement Levels

- Provide meaning and consequence to the scale
- Build consensus regarding the requisite knowledge and skills that students are required to master
- Shine a spotlight on inequities in the education system
- Support state and federal accountability systems
- Support state and federal retention laws (e.g. grade and ELL)
Achievement Levels

- **Involves Expert Judgement**
  - Educators who understand the grade level content
  - Educators that are knowledgeable re: the diversity of students who take the test
  - Educators who are knowledgeable re: grades either above or below

- **Build consensus regarding the content that is assessed**
Achievement Levels

- Moderated by:
  - Test format
  - Politics
  - Consequences of false positives and false negatives
  - Cost and logistics of true positives and true negatives

- Resources available
  - For instructional days
  - Professional learning
  - Textbooks and curriculum supports

- Test administration windows
Achievement Levels
Wayne Camara
### Achievement Levels

<table>
<thead>
<tr>
<th>EDUCATION</th>
<th>CERTIFICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EXCEEDS - What you can do</strong></td>
<td><strong>PASS</strong></td>
</tr>
<tr>
<td><strong>MEETS - What you can do</strong></td>
<td><strong>FAIL</strong></td>
</tr>
<tr>
<td><strong>APPROACHES - What you can do</strong></td>
<td></td>
</tr>
<tr>
<td><strong>PARIALLY MEETS - What you can do</strong></td>
<td></td>
</tr>
</tbody>
</table>

#### EXCEEDS - MATH ALGEBRA
Algebraically solves linear equations, linear inequalities and quadratics in one variable (at complexity appropriate to the course), including those with coefficients represented by letters. Utilizes structure and rewriting as strategies for solving.

#### MEETS - MATH ALGEBRA
Algebraically solves linear equations, linear inequalities and quadratics in one variable (at complexity appropriate to the course).
Can Credentialing Tests Provide Better Achievement Level Descriptors to Test Takers and Test Users?

- Maybe Yes, Maybe No?
- Legal Risks
- Ability to match candidate data to outcome data
- Ability to match candidate data to prior experience, accomplishments, and education
- Will it matter, is it important to users and candidates?
- Do we focus on job performance, KSAs, other outcomes?

Example:

Students have obtained the preparation in math such that they have at least a 70% chance of enrolling, without remediation, in entry level college credit and career training courses in math.
Example ALDs for Credentialing

*Candidates who have met expectations:*

- Are 3x as likely to avoid having any safety violations during their first year of employment (railroad engineer).
- Are 3x less likely to have a charge of malpractice during their first 5 years (surgeon, criminal attorney)
- Are 3x as likely to complete mandatory PD required to keep their license to practice (any field)
- Are 3x as likely to be employed within six months
- Have worked an average of 2+ yrs more than those who didn’t pass (CPA, nurse)
- Have completed an average of 20 more credit hours in STEM courses than those who didn’t pass (nurse)
- Have attended on average 2x PD workshops (IT)
What is alignment?
What is alignment?

“the degree to which expectations and assessments are in agreement”

(Webb, 1997, p. 3)
What is alignment?

“the degree to which expectations and assessments are in agreement”

(Webb, 1997, p. 3)

“the degree to which the content and cognitive demands of test questions match targeted content and cognitive demands described in the test specifications”

(AERA, APA, & NCME, 2014, p. 216)
Two Reasons to Conduct Alignment Studies
Two Reasons to Conduct Alignment Studies

1. Content standards only define the domain; provide evidence of intention
Two Reasons to Conduct Alignment Studies

1. Content standards only define the domain; provide evidence of *intention*
2. Alignment analyses provide information on test construction success, validity evidence
Some Alignment Methods
Some Alignment Methods

- Surveys of Enacted Curriculum
  (SEC; Porter & Smithson, 2001)

- Achieve Method
  (Rothman, Slattery, Vranek, & Resnick, 2002)

- Webb Method
  (Webb, 1997)
Some Alignment Methods

- **Surveys of Enacted Curriculum**  
  (SEC; Porter & Smithson, 2001)

- **Achieve Method**  
  (Rothman, Slattery, Vranek, & Resnick, 2002)

- **Webb Method**  
  (Webb, 1997)

- **Generalized Assessment Alignment Tool (GAAT)**  
  (Cizek & Kosh, 2016)
The Generalized Assessment Alignment Tool (GAAT)
The Generalized Assessment Alignment Tool (GAAT)

- Curriculum Coverage (CC1)
The Generalized Assessment Alignment Tool (GAAT)

- Curriculum Coverage (CC1)
- Construct Comprehensiveness (CC2)
The Generalized Assessment Alignment Tool (GAAT)

- Curriculum Coverage (CC1)
- Construct Comprehensiveness (CC2)
- Content Concentration (CC3)
The Generalized Assessment Alignment Tool (GAAT)

- Curriculum Coverage (CC1)
- Construct Comprehensiveness (CC2)
- Content Concentration (CC3)
- Cognitive Complexity-Absolute (CC4a)
The Generalized Assessment Alignment Tool (GAAT)

- Curriculum Coverage (CC1)
- Construct Comprehensiveness (CC2)
- Content Concentration (CC3)
- Cognitive Complexity-Absolute (CC4a)
- Cognitive Complexity-Relative (CC4b)


**Educational Alignment**

1. Create Content Standards (e.g., 5th grade math)
2. Identify content experts
3. Determine the organization of the Standards (strands, major areas, areas of practice)
4. Identify what students should know and be able to do across grades - developmental or learning progression
5. Determine eligible or assessable content (measurable, relevance) at grade level
6. Prioritize in developing a content blueprint for the assessment. Based on major facets - content coverage, breadth, depth of knowledge
7. Focus on what should be taught tomorrow (future)

**Occupational Job Analysis**

1. Develop list of tasks required to perform a job successfully (e.g., teacher aide)
2. Identify subject-matter-experts (SMEs)
3. Identify the most critical tasks (SME importance ratings)
4. Identify critical competencies or KSAs
5. Link KSAs to critical tasks
6. Determine the importance of competencies (or KSAs) based on linkage to tasks.
7. Determine it competencies will be assessed (or developed after selection). Map test blueprint to competency profile.
8. Focus on what competencies are needed today (current-past)
Educational Alignment

1. Change how teachers teach, not reflect current practice
2. Standards finalized through consensus at the end of the process
3. Policymakers or an external framework may drive the outcome (e.g., more rigor, what skills are taught in what grades)
4. Panels of experts generally are responsible for the entire process. Rarely is there any cross-validation by independent raters
5. No statistical evidence of ratings or judgments is used
6. Standards become the basis for an assessment design and the primary evidence to support a validity argument so how they are developed, who develops them, and whether they reflect reality or aspiration is important

Occupational Job Analysis

1. Reflect current performance, not change how incumbents perform the job
2. Competencies and tasks determined through statistical criteria and ratings
3. Job descriptions, current practices will drive the outcome and may not reflect best practices
4. Initial competencies, KSAs may be established by a small panel, but these lists are expanded and revised based on ratings collected from diverse set of SMEs (incumbents, supervisors, experts) based on evidence of convergence
5. Statistical evidence is available to examine issues such as agreement, reliability, bias. Job analysis generally conducted to reflect reality and content validity
6. Job analysis often become the basis for selection, evaluation and training processes, as well as the primary evidence for a validity argument so how they are developed, who develops them and whether they reflect actual job requirements is important
Issues with Educational Content Standards

- Standards-
  - Do they reflect leading-edge/innovation or consensus/practice?
  - What is the format, level of specificity, prescription?

- SMEs -
  - Do we balance representation or stress professional (expert) participation?
  - Do we involve higher ed faculty (HS)? What level of non-content expertise is involved (parent, employer, counselor)

- Outcome -
  - Often predetermined by policymakers, state leadership, or external perspectives most attractive, rather than focus on curriculum.
  - Standards have limited validity evidence; focus often on implementation rather than effectiveness.
Recommendations for Improving Educational Standards

- Begin with current practices and needs
- Broaden beyond an SME panel - curriculum surveys, best practices
- Use ratings from teachers at grade level, above level and below level
- Use statistical methods to determine convergence, agreement and to identify outliers
- Consider ways to validate standards - predictive validity studies showing if students possessing skill actually succeed in course or on assessment; concurrent validity - do successful students possess skill?
Standards

1.9 When a validation rests in part on the opinions or decisions of expert judges, observers, or raters, procedures for selecting such experts and for eliciting judgments or ratings should be fully described. The qualifications and experience of the judges should be presented. The description of procedures should include any training and instructions provided, should indicate whether participants reached their decisions independently, and should report the level of agreement reached. If participants interacted with one another or exchanged information, the procedures through which they may have influenced one another should be set forth.

11.13 The content domain to be covered by a credentialing test should be defined clearly and justified in terms of the importance of the content for credential-worthy performance in an occupation or profession. A rationale and evidence should be provided to support the claim that the knowledge or skills being assessed are required for credential-worthy performance in that occupation and are consistent with the purpose for which the credentialing program was instituted.
Test Length

Tony Alpert
Test Length & Related Considerations

- **Purpose**
  - Support improvements in teaching and learning
  - Signal high quality instruction
  - Shine a spotlight on inequities in the education system
  - Address federal and state reporting and accountability laws
  - Requires that we measure the breadth and depth of the content standards
Test Length & Related Considerations

- Consequences of the test
  - Types of decisions based on the results
  - Availability of benefits and/or opportunities

- Test taker attributes
  - Age
  - Significant Cognitive Disabilities

- Legal and political constraints
  - Statutes that restrict testing time
  - Statutes that require other tests

- Cost

- Technology limitations
Test Length & Related Considerations

Test Format
- Evidence centered design (measure what we intend to measure)
- Accessible
- Generalize results to real world situations
- Concerns around machine scoring
Test Length
Liberty J. Munson
Test Length & Related Considerations

- IT certification
  - Voluntary certification
  - Recertification requirements
  - Pass/fail decisions are key, actual score less so

- Test Format
  - Computer-based vs. Paper-based
  - Multiple choice, interactive, performance based
  - Adaptive, dynamic, static delivery
Test Length & Related Considerations

- Test Length
  - Seat time and exam scheduling with EDPs
    - Candidates want shorter exams
    - Customer satisfaction with experience (NSAT)
  - Reliability and validity
- Item writing
  - Follow blueprint
  - Difficulty and discrimination geared towards distinguishing borderline candidates
Test Security
Gregory J. Cizek
Test security is even more consequential in a consortium environment
Context

- Test security is even more consequential in a consortium environment
- NextGen ELA, Science, Math tasks are even more expensive to develop
Context

- Test security is even more consequential in a consortium environment
- NextGen ELA, Science, Math tasks are even more expensive to develop
- Many reasonable policies are in place...
Context

- Test security is even more consequential in a consortium environment
- NextGen ELA, Science, Math tasks are even more expensive to develop
- Many reasonable policies are in place...
- ...but, policy is not enough
Test security is even more consequential in a consortium environment

NextGen ELA, Science, Math tasks are even more expensive to develop

Many reasonable policies are in place…

…but, policy is not enough

More rigorous detection, investigation, responses necessary
NextGen Test Security
NextGen Test Security

- **NextGen Conceptualization**: *Security not separate, but nested under validity*
The *Standards for Educational and Psychological Testing* should have four sources of validity evidence.
The *Standards for Educational and Psychological Testing* should have four sources of validity evidence.

Evidence based on...

1. Test Content
The *Standards for Educational and Psychological Testing* should have four sources of validity evidence.

Evidence based on...

1. Test Content
2. Response Process
The *Standards for Educational and Psychological Testing* should have four sources of validity evidence.

Evidence based on...

1. Test Content
2. Response Process
3. Hypothesized Relationships among Variables
The *Standards for Educational and Psychological Testing* should have four sources of validity evidence.

Evidence based on...

1. Test Content
2. Response Process
3. Hypothesized Relationships among Variables
4. Test Development and Administration Procedures
NextGen Test Security

- NextGen Conceptualization: Security not separate, but nested under validity

- **NextGen Detection**: Many appropriate methods exist (see for example Cizek & Wollack Handbook)
NextGen Test Security

- NextGen Conceptualization: Security not separate, but nested under validity

- NextGen Detection: Many appropriate methods exist (see for example Cizek & Wollack Handbook)

- NextGen Paradigm: Sufficient statistics
In the 1990s...

“It would seem prudent that the weight of statistical evidence only be brought to bear when some other circumstances—that is, a trigger—provide a reason for flagging cases for subsequent statistical analysis.”

(Cizek, 1999, p. 142)
“In some cases, absent compelling counter-evidence, rigorous statistical methods yielding highly improbable findings might alone be sufficient to conclude that an examinee’s test score should not be considered a valid representation of his or her knowledge or ability.”  

Cizek & Wollack, 2017, p. ii)
NextGen Test Security

NextGen Conceptualization: Security not separate, but nested under validity

NextGen Detection: Many appropriate methods exist (see for example Cizek & Wollack Handbook)

NextGen Paradigm: Sufficient statistics

NextGen Models:
NextGen Test Security

- **NextGen Conceptualization:** Security not separate, but nested under validity
- **NextGen Detection:** *Many appropriate methods exist* (see for example Cizek & Wollack *Handbook*)
- **NextGen Paradigm:** *Sufficient statistics*
- **NextGen Models:** In some Lic/Cert areas, “Passing” requires meeting two criteria:
NextGen Test Security

- NextGen Conceptualization: Security not separate, but nested under validity
- NextGen Detection: Many appropriate methods exist (see for example Cizek & Wollack Handbook)
- NextGen Paradigm: Sufficient statistics

- NextGen Models: In some Lic/Cert areas, “Passing” requires meeting two criteria:
  1. Cut score threshold
NextGen Test Security

NextGen Conceptualization: Security not separate, but nested under validity

NextGen Detection: Many appropriate methods exist (see for example Cizek & Wollack Handbook)

NextGen Paradigm: Sufficient statistics

NextGen Models: In some Lic/Cert areas, “Passing” requires meeting two criteria:
1. Cut score threshold
2. Score validity threshold
Test Security
Liberty J. Munson
Test Security

- Arguably the most significant threat against credential relevance and reliability

- Evolving threats from “cheating” industry
  - Braindump sites, proxy testing, traditional cheating, falsified score reports, fraudulent claims yet to be discovered
  - Hidden cameras, spy devices
What is cheating?

Collusion

Any activity that enables unqualified candidates to pass exams

Falsified Score reports

Braindumps

Proxy Testing
Scope of the Problem

IT Industry-wide Problem

As technology grows, so does fraud

Hiring Manager’s Concerns

Candidate Concerns

Cost to Credentialing Body

- Costs associated with new content development
- Program reputation and integrity
Certification Approach

Protecting Integrity Prior to Testing Event

Prevents current unqualified candidates from being certified

Protecting Integrity After Testing Event

Prevents future unqualified candidates from being certified
## Approach to Pillars

<table>
<thead>
<tr>
<th>Education</th>
<th>Protection</th>
<th>Investigation</th>
<th>Enforcement</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Website</td>
<td>• NDA</td>
<td>• Leads</td>
<td>• Candidate Bans</td>
</tr>
<tr>
<td>• Program Reports</td>
<td>• Policies</td>
<td>• Investigative Team</td>
<td>• Test Site Closures</td>
</tr>
<tr>
<td>• Press Releases</td>
<td>• Secure Testing Environment</td>
<td>• Test Purchases</td>
<td>• C&amp;Ds and Take Down Notices</td>
</tr>
<tr>
<td>• Stories/blogs</td>
<td>• Delivery (dynamic, continuous publication, etc.)</td>
<td>• Secret Shops</td>
<td>• Lawsuits</td>
</tr>
<tr>
<td>• Blogs</td>
<td>• Geo Blocking</td>
<td>• Internet Monitoring</td>
<td></td>
</tr>
<tr>
<td>• Presentations</td>
<td></td>
<td>• Auction Monitoring</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• OOC</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Data Forensics</td>
<td></td>
</tr>
</tbody>
</table>
Questions?
Thank you!!!!

- Tony.alpert@smarterbalanced.org
- Wayne.camara@act.org
- Cizek@unc.edu
- Liberty.munson@Microsoft.com
- Jamie.mulkey@Caveon.com