Online Assessment: Panacea, Pandora’s Box, or Somewhere In Between? Best Practices, Thoughts, and Current Challenges with Un-proctored and Proctored Online Testing

IPMAAC
St. Louis, MO
June 11, 2007
Representative Perspectives

- David Hamill (Session Chair)
- Sebina Netto (Vendor/Public Sector)
  - JobAps
- Jim Beaty, Ph.D. (Vendor/Private & Public Sector)
  - PreVisor
- Jeff Feuquay, Ph.D., Esq. (Legal)
  - Attorney & Psychologist
- Eric Palmer (Public Sector User)
  - City of Ft. Worth
- Michael Blair (Private Sector User)
  - Embarq

June 11, 2006
Session Format

- Individual Panel Members
- Panel Discussion of Hot Topics in UIT
- Audience Q&A
- Final Thoughts
Attitudes Towards UIT

- Unproctored Internet testing (UIT) for has too many pitfalls and drawbacks to be used in for employment purposes.
  1. Strongly Disagree
  2. Moderately Disagree
  3. Neutral
  4. Moderately Agree
  5. Strongly Agree
Types of Unproctored Internet Testing

- Open
- Invitational
- Supervised
Sabina Netto

Directs Customer Implementations at JobAps

Prior to joining JobAps, she was a Supervising HR Analyst at the County of Santa Barbara, California.
Our Customers are the Experts - Interviews with our Customers

- State of California - Daisy McKenzie
- State of Oklahoma - Hank Batty
- State of Delaware - Cheryl Rice
- Hillsborough Co., Florida - Kurt Wilkening
Oscar Jackson - State of Oklahoma

Director OPM, President of IPMA

Warning:

“Bring a roll of chicken wire to protect the panel from the tomatoes that your audience might throw.”

June 11, 2006
Daisy McKenzie - **State of California**
Manager of Exam Services

- Success- for professional classes that have license, degree, certification (already tested)
  1. Self-rated online MQ screen
  2. Followed by self-rated online T&E
Daisy McKenzie - State of California
Manager of Exam Services

- Failure - Staff Services Analyst

1. Self-rated online T&E - criterion evaluation of test - Failure (estimate 50% falsification)

2. Redid Job Analysis/Validation Study → Online proctored M/C Test (1% dismissed for cheating)
Daisy McKenzie - State of California
Manager of Exam Services

- Biggest Surprise
  “How many resources we saved once we figured out how to automate”
  “We calculate the effort is reduced by 90%”
Kurt Wilkening - Hillsborough Co, Florida

Personnel Research Manager

- Would like to eventually have a proctored online test environment.
- Does not believe that non-proctored online tests will work due to cheating - unless used as a screen with follow up proctored testing.
Kurt Wilkening - Hillsborough Co, Florida

Personnel Research Manager

- How about an Online Non-Proctored Power Test (Timed Test)?

It may help eliminate cheating, but we would not want to exclude people who may think more slowly.
All classes utilize a self-scoring and self-screening supplemental questionnaire that assesses competencies using different rating scales.

Legislative change accompanied automation project - refer top 30 scores instead of 10
Cheryl Rice - **State of Delaware**

Automation Project Manager

- **Results** - number of applications has increased 3 to 4 times
- **Hiring Agencies** love it
- **More metrics** to come

June 11, 2006
Hank Batty - State of Oklahoma

Deputy Director, OPM

- Electronic testing since 2003, always provide a paper option - 80-90% online.
- “Biggest surprise - how quickly applicants took to it.”
Hank Batty - State of Oklahoma

Deputy Director OPM

- “I know electronic testing is more secure than paper and pencil.”
- “Big savings from non-proctored testing, but people will cheat.”
Hank Batty - State of Oklahoma
Deputy Director OPM

“Smart People make better employees than people who are not smart. If the test discerns that - it is helpful. We can teach skills.”

June 11, 2006
Hank Batty - State of Oklahoma
Deputy Director OPM

❖ “Online Testing is not a panacea, but if used as intended, it is cool stuff and can help you.”
Jim Beaty, Ph.D.
PreVisor

- Chief Scientist
- Oversee R & D at PreVisor
- 10 years experience doing personnel selection
- 7 years experience developing unproctored internet tests
PreVisor’s Top 50 clients

- 40% Unproctored
- 60% Proctored Only
UIT norm trends – Entry

<table>
<thead>
<tr>
<th>Job</th>
<th>N/d</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006*</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Labor</td>
<td>&gt; 9,000</td>
<td>+.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales</td>
<td>&gt; 5,000</td>
<td>+.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales</td>
<td>&gt; 2,000</td>
<td>-.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Call Center</td>
<td>&gt; 2,000</td>
<td>+.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Customer Service</td>
<td>&gt; 8,000</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Jan-Feb 2006

June 11, 2006  IPMAAC
## UIT norm trends – Professional

<table>
<thead>
<tr>
<th>Job</th>
<th>N/ d</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td>&gt; 30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College Recruits</td>
<td>&gt; 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales</td>
<td>&gt; 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>+.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manager</td>
<td>&gt; 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>+.06</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Jan-Feb 2006
# Recent UIT Validity studies

(Competency/Solution Scores)

<table>
<thead>
<tr>
<th>Metric</th>
<th>N</th>
<th>Corrected r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales production</td>
<td>434</td>
<td>.17/.13</td>
</tr>
<tr>
<td>Retail sales metric</td>
<td>480</td>
<td>.17/.10</td>
</tr>
<tr>
<td>Call Center sales metric</td>
<td>662</td>
<td>.26/.22</td>
</tr>
<tr>
<td>Collections metric</td>
<td>232</td>
<td>.28/.16</td>
</tr>
<tr>
<td>Call Center - Call quality rating</td>
<td>255</td>
<td>.16/.14</td>
</tr>
<tr>
<td>120-Day Retention – Call Center</td>
<td>429</td>
<td>.18/.16</td>
</tr>
</tbody>
</table>

June 11, 2006

IPMAAC
Cheating Risk

- Easily identifiable testing process
- Testing process is not easily identifiable
- Severe penalty for cheating

Low Stakes
- entry to mid level
- test not obvious hurdle

High Stakes
- exec level
- test is obvious hurdle

Cheating risk:
- High
- Low

June 11, 2006

IPMAAC
If you have ever been to a certification exam, you know how hard it is, to pass such exams. You may spend big money in books & practicals, training camps, but still cannot be sure of passing. We are here to put an end to all your worries. Use our Cheat-test study guides and you’ll not need anything to pass the toughest of exams.
Defending against cheating

- Document the business case for implementing unproctored testing
- Design your system with a single point of entry into the online hiring process
- Maintain current recruiting efforts
- Follow unproctored tests with a proctored assessment where applicant identity is confirmed
- Warn applicants not to cheat
- Monitor and validate test scores over time.
Internet Testing Experience

- Prosecute & defend selection lawsuits
- Decades of selection experience
- Extensive review of internet testing research and articles
Feuquay Issues & Premises

- Test purpose, type & administration conditions
  - Proctored ↔ Unproctored
  - Low ↔ High Stakes
  - Phase in selection process
- Ethics issues, e.g., Principle 9.09(a), APA
- All the usual legal issues (Hey, they’re tests.)

June 11, 2006
Eric Palmer, SPHR
City of Fort Worth Texas

- 7 years in public sector HR, much of it spent making the business case for investing in effective HR practices
- Ongoing interest in developing selection innovations for the public sector
- Experience working at the developmental stages of online delivery systems, including at City of Fort Worth Texas, Santa Barbara City College, and Santa Barbara County.
- Worked with Jobaps during early stages
In Defense of High Stakes UIT

Pro and Con: where do the sides agree?

- Difficulties controlling test environ form a strong *prima facie* case against high-stakes UIT
- Strong argument in favor is needed before high-stakes UIT warrants serious discussion

Where do we disagree?

- There *is* such a favorable argument: the HUGE value-added potential of high stakes UIT
- Technological v. adaptive changes
  - The obstacles are technological and can be overcome
  - The benefits are adaptive, and have the potential to revolutionize how many of us approach talent acquisition
What Problem can UIT Address?

**Applicant Pipeline Congestion:** applicant flow exceeds capacity; inability to process large numbers both efficiently and effectively

- Panel interview is usually the primary selection tool, but can only interview a few
- Quandary: how to get from large potential applicant pool to a few interviewees?

**Common maladaptive “solutions”**

- Minimize applicant pool by minimizing ads and time open
- Apply an efficient but poorly validated screening tool to the applicant pool

**Solution: high-stakes UIT**

June 11, 2006
The Business Case for UIT

- Replace poorly validated employment standards with highly validated tests of critical skills
- Merge application, MQs, and testing
- UIT with a validated cognitive test can screen huge numbers of applicants for true critical skills
- This enables a strategy of “maximizing the applicant pool”
  - Maximize the applicant pool
  - Screen in high performers based on true critical skills
  - Maximize the eligible list
  - Maximize hiring of high performers
  - Develop a high performing organization

June 11, 2006 IPMAAC
Obstacles to UIT

- **Most cheating doesn’t matter!**
  - Cheating only matters if it results in someone on the eligible list who would not be there otherwise.
  - Below the eligibility list cutoff, score movement as a result of cheating doesn’t matter.
  - Most cheating does not result in significant score movement.
  - It is unusual for someone in the top 10% to be there due to cheating.

- **Cognitive v. non-cognitive UIT**
  - Non-cognitive is safer, but you lose one of the key benefits: a rigorously valid and predictive exam that can be efficiently administered to thousands.
  - Use cognitive UIT to *screen in* top performers rather than *screen out* poor applicants.
# HOW DOES UIT COMPARE?

<table>
<thead>
<tr>
<th>DAY</th>
<th>UIT SELECTION PROCESS</th>
<th>TRADITIONAL SELECTION PROCESS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 Recruitment opens</td>
</tr>
<tr>
<td>day 1-14</td>
<td><strong>Recruitment opens</strong></td>
<td><strong>Recruitment opens</strong></td>
</tr>
<tr>
<td>day 1-14</td>
<td><strong>Online applications accepted</strong></td>
<td><strong>Online applications accepted</strong></td>
</tr>
<tr>
<td></td>
<td>Applicants complete a unique profile in the system that acts as a gateway to the online examinations</td>
<td>Applicants complete a T&amp;E and supplemental questionnaire</td>
</tr>
<tr>
<td></td>
<td>Applicants complete the initial (unproctored) online cognitive abilities exam (merge MQ screen &amp; selection exam)</td>
<td></td>
</tr>
<tr>
<td>End of 14</td>
<td><strong>Recruitment closes</strong></td>
<td><strong>Recruitment closes</strong></td>
</tr>
<tr>
<td>15</td>
<td>Scores from initial exam are standardized</td>
<td>15-20 Applications screened for minimum employment standards</td>
</tr>
<tr>
<td></td>
<td>&quot;Screen-in&quot; score established at 10th lowest score</td>
<td>21-26 Resumes and supplemental questionnaires are reviewed to identify the 10 most qualified applicants</td>
</tr>
<tr>
<td></td>
<td>Top 10 scoring candidates invited to proctored exam</td>
<td>27 Top 5-10 candidates invited to panel interviews</td>
</tr>
<tr>
<td>22</td>
<td>Proctored exam administered</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exams scored, scores standardized</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Applicants scoring below initial &quot;screen-in&quot; score are eliminated</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Remaining candidates ranked by score</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Top 5-10 candidates invited to panel interviews</td>
<td></td>
</tr>
<tr>
<td><strong>8 Days from close of recruitment to invitation to panel interviews</strong></td>
<td><strong>13 days from close of recruitment to invitation to panel interviews</strong></td>
<td></td>
</tr>
</tbody>
</table>

June 11, 2006

IPMAAC
Michael Blair, EMBARQ

- Selection & Assessment Program Manager
- Oversees Entry-Level and Promotional Assessments for EMBARQ
- 15 years experience in personnel selection & assessment
  - Public & private sector
- 10 years experience in high stakes testing
EMBARQ

Who is EMBARQ?
- Total communication company
  - Local, long distance, wireless, high-speed internet, digital video
  - 20,000 employees serving customers in 18 states
    - 4.8 million households and 500,000 businesses

Online Assessment at EMBARQ
- Entry-level positions in call centers, retail stores, business sales, and other customer facing positions
- Front-line supervisors in our Consumer Division
- Screening and selection measures in both proctored and unproctored settings
- Internet-based testing is expected to exceed 10,000 applicants for calendar year 2007
Candidates as Stakeholders

- Candidates want:
  - A clear relationship between test content and job content
  - *Fairness and accuracy* in test scores
  - Processes that are actually needed
  - *Ease of use* and less demanding content

- Candidates prefer online assessments
  - *Preferred* over more traditional based processes
    - Frequency of administration
    - More timely or immediate feedback
    - Require less testing time

- The candidate experience is important
  - Recruitment
  - Job Satisfaction
  - Organizational Commitment
  - Job Performance
  - Turnover
Incumbents, Supervisors, & Managers As Stakeholders

- This group of stakeholders want:
  - Fairness and accuracy
  - Processes that are related to the job
  - Processes that identify good employees and good co-workers
    - Performance
    - Teamwork
    - Fit

- In addition, supervisors and managers want:
  - Ease of use
  - Results that are consistent across types of assessments
    - E.g., online test scores that align with interview results
  - Results that align with expectations of incumbent ability
  - The perfect process

June 11, 2006  IPMAAC
Decision Makers As Stakeholders

- Decision Makers Expect:
  - Fairness
  - Processes that are aligned with business goals
    - Diversity
    - Customer Service
    - Innovation
  - Return on Investment (ROI)
    - Online assessments offer bang for the buck
Hot Topics in UIT

- Cheating
- Security of Test Content
- Identity Concerns
- Purposes of Assessment
- Validity and Reliability
- Policy Issues
- Over Automation
- Legal Pitfalls
- Adverse Impact
- Hardware & Software Standardization