

ICE Digest: 2015 Q3 Edition

Message from the Executive Director

[ICE Expands its Reach in 2015](#)

By Denise Roosendaal, CAE

Features

[6 Organizations Collaborate on Survey to Determine How Certification Enhances Salary Potential](#)

By Jeffrey H. Greenwald, PE, CAE

[Overcoming the Threats and Risks of Testing Internationally](#)

By Dennis Maynes and Aimée Rhodes

[Putting the Principles of Assessment Engineering into Practice — Impacts for the Design and Analysis of Job Analysis Surveys](#)

By Matthew Schultz and John Mattar

[Test Security: A Look Behind the Scenes](#)

By Joy L. Matthews-López, PhD; Ada Woo, PhD; Alan J. Thiemann, JD; Paul E. Jones, PhD; and Jennifer Gallagher

[5 Questions to Ask Before Choosing a Remote Proctoring Provider](#)

By Jarrod Morgan

[A Psychometrician's Musings on Recertification: What it Means, What You're Really Assessing and How to Ensure Fairness](#)

By Liberty Munson, PhD

Special Thanks to the Industry Partners Who Made the 2015 Q3 Edition Possible



ICE Expands its Reach in 2015

By Denise Roosendaal, CAE



ICE is reaching beyond its traditional membership borders and connecting with professionals to advance the understanding of credentialing and promote best practices.

ICE recently presented a session at the ASAE Annual Conference in Detroit. The topic was “What’s New in the Business of Certification” and was presented by ICE members Karen Plaus, PhD, CRNA, FAAN, CAE, and Maggie Collins, PhD. The topic allowed ICE to showcase the 2015

Business of Certification report released by the ICE Research and Development Committee. The session attracted about 100 attendees who engaged with many thoughtful questions about research and certification best practices. In advance of the conference, the content from the research was turned into a brief article highlighted in the electronic version of *Associations Now*.

ICE’s director of accreditation, Anjali Weber, has presented several times to the Association Trends audience. Her latest presentation was on Sept. 18 on the topic of international certification.

In April, ICE members presented at and attended the American Legion Summit where the topic of private sector credentialing for active duty military and veterans was discussed. This summit was an opportunity for policymakers, military leaders and private sector industry leaders to convene on ways to create more pathways to assist the transition from active duty to private sector.

The ICE social media plan has opened up many opportunities to highlight credentialing best practices. ICE has also been participating in ASAE’s listserv on a variety of credentialing topics.

The credentialing chapter in ASAE’s Professional Practices in Association Management, released in early 2015, was written by Anjali Weber and me. This chapter will hopefully inform many professionals who might be just learning about credentialing.

Further exposure occurred when I was recently interviewed by *Bottom Line Briefing*, an electronic newsletter sent to association CEOs. It highlighted various aspects of credentialing and the topic of professionalization of professions and occupations.

ICE is also re-engaging in the government affairs arena and monitoring legislation and regulation that impacts credentialing. Recently, David Swankin attended an invitation-only session at the White House to discuss the state of licensure and occupational trends from around the country. He eloquently brought certification to the conversation as an alternative to licensure in some cases.

While these are just a few examples from this year, ICE will continue expanding its reach in order to expand its message of certification best practices and the emergence of certificate programs.

6 Organizations Collaborate on Survey to Determine How Certification Enhances Salary Potential

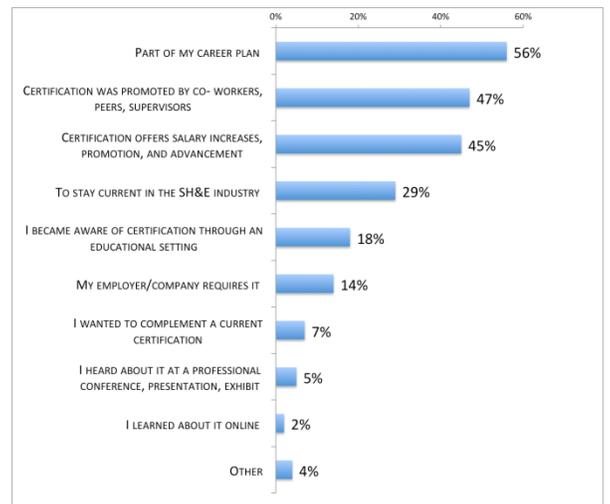
By Jeffrey H. Greenwald, PE, CAE

As the business world becomes more complex and specialized, the demand for individuals with industry-recognized credentials will continue to intensify. Corporate governance reforms, an increasingly global business environment and rapidly changing technologies have fueled the need for professionals with advanced skills and up-to-date knowledge.

Companies value professionals who proactively seek opportunities to develop their skills and expertise. One of the best ways for individuals to demonstrate their commitment to career growth is by pursuing recognized certifications. The eligibility requirements are tough. Preparation can be a lot of work. The examination can be hard. Everyone has a different reason for pursuing a voluntary certification (e.g., peer recognition, personal satisfaction, job placement or advancement, market value, higher salaries). For some, the rewards are immediate and outward; for others, the rewards are intrinsic.

Six organizations within the environment, health and safety professions collaborated to conduct a salary survey to find out how certification enhances salary potential. The six organizations were the American Board of Industrial Hygiene (ABIH); Alliance of Hazardous Materials Professionals (AHMP); American Industrial Hygiene Association (AIHA); American Society of Safety Engineers (ASSE); the Board of Certified Safety Professionals (BCSP); and the Institute of Hazardous Materials Management (IHMM). Data for the survey was collected from March 26, 2015, through April 7, 2015, and the survey was closed for tabulation with 9,258 responses. Each organization was requested to provide certificants/member lists and, through a series of teleconferences, oversaw the project. Staff management of the project was provided by BCSP.

The results were favorable. Earning accredited credentials can provide numerous benefits, including improved career prospects and enhanced earning power. According to the salary survey, the results indicate that the median annual base salary of individuals holding at least one credential is \$98,000, about 18 percent greater than the median annual base salary of survey respondents who held no certification: \$83,000. Furthermore, when years of experience are taken into account, salaries are approximately 12 percent higher for certified professional holders with five to 14 years of experience than a professional with the same experience not holding a credential. Finally, when salaried respondents were asked why they pursued certification, 56 percent said it was part of their career plan.



The results of the salary survey can be found at the collaborating organizations' websites.



Overcoming the Threats and Risks of Testing Internationally

By Dennis Maynes and Aimée Rhodes

International expansion of credentialing programs presents exciting opportunities along with significant threats and risks. Exam and program security are critical for delivering good quality growth. Whenever a credentialing program enters a new country, specific security measures should be considered to help preserve the integrity of exams and to generate increased confidence in the quality of test administration. In order to manage threats found in other countries, programs need to quantify operational risks and implement processes that address vulnerabilities, protect against attacks and mitigate risks.

Vulnerabilities are weaknesses in a testing program that can be exploited, or failures that can occur naturally resulting in loss. **Threats** are potential actions or occurrences resulting from one or more vulnerabilities. **Attacks** are the realization of threats that may or may not be successful. **Risks** are possible losses that result from threats and attacks.

Because exam security problems are regularly reported by international media, credentialing programs may be reluctant to embrace international expansion. For example, Jennifer Semko, writing for the *Federation Forum*, described a Philippines-based test preparation company that received and distributed “a comprehensive compilation of test questions recalled by past test takers” (Semko, fall 2007, “The Story behind the NPTE Score Invalidations”). Many other programs have had similar experiences. On the other hand, the opportunities and advantages of international testing cannot and should not be denied. Our purpose in writing this is to share perspectives and key learnings that will aid your program as you consider international expansion.

Vulnerabilities of Testing Internationally

Whether you are testing at home or abroad, your program will be subject to most of the same vulnerabilities:

- Individuals can harvest (steal) test items in the same ways.
- The physical security of testing centers may be breached in the same way.
- Exams can still be diverted in transit.

International exam security challenges usually result from increased likelihood that vulnerabilities already present will be exploited. Additionally, you may encounter new vulnerabilities. Threats and attacks are more likely to occur when testing in other countries because increased opportunities and motivations to attack are present in more diverse candidate populations (e.g., due to cultural values, varying socioeconomic conditions, ideological differences).

Threats of Testing Internationally

The increased likelihood of exam security threats is driven by three main factors: 1) new vulnerabilities that arise from geographical distances and sociopolitical environments; 2)

cultural differences (such as the social desirability of cheating or the acceptance of bribing officials); and 3) high stakes from passing exams.

As an example of increased threats due to high stakes, consider the prevalence of “qiangshou” (literally “gunman”) in China. On April 19, 2014, the *Shanghai Daily* published a cartoon under the headline: “Hired ‘guns’ in demand to pass tests.” The article described four Chinese individuals who had taken exams for others in more than 25 countries before being caught in South Korea. In some areas of the world, gunmen or proxy test takers are in high demand because passing the test could be the difference between a promising and a bleak future. Especially pervasive in some countries are electronic devices that enable cheating. Some technically advanced devices use Bluetooth to link audio and video with cellphones. These devices may be hidden almost anywhere (e.g., as a piece of jewelry, in a tie or in glasses). On May 31, 2011, CBC News in British Columbia reported that two men used electronic devices to transmit items from the Medical College Admissions Test to tutors. The Association of Medical Colleges “told police the incident cost them more than \$200,000 because the compromised test had to be discontinued and replaced.”

We recommend your organization evaluate the following elements to understand threats in a particular country before you test there:

1. Overall political landscape
 - How stable is the government?
 - Is civil unrest common?
 - Is regional terrorism a concern?
2. Firsthand information from other testing programs
 - Have any staff members been threatened or assaulted?
 - What is the experience with fraudulent financial transactions?
 - Have impersonators interfered with the organization’s ability to register test takers?
 - How widespread is cheating on tests?
3. Travel warnings and advisories that may have been issued for foreign nationals
 - Have travel advisories been issued for the country?
 - Does your organization prohibit travel to the country?
4. Issues with transmission of exam material
 - Can secure exam content be transferred through customs without being compromised?
 - Can you electronically transmit data (exam content, biometrics, CCTV video) to and from the country without interference?

Risks of Testing Internationally

When testing internationally, some losses have higher probabilities and greater harm than others. For example, some Central African countries are dealing with terrorist organizations that have been known to kidnap/kill foreigners.

When exam security attacks are directed by organizations (e.g., Internet-enabled pirates, well-funded commercial enterprises or time-zone cheating syndicates), the probability of loss goes up. Risks are greater because the organizations have more resources for pursuing the attack. Many of these organizations may exploit international testing vulnerabilities (e.g., when the same forms are administered in multiple time zones). An organized group of item harvesters can steal an entire exam efficiently and quickly.

Also, remedies used successfully by U.S.-based credentialing programs may not work in other countries. For example, when foreign websites that distribute harvested exam content disregard Digital Millennium Copyright Act takedown notices, the entire investment in developing the exam may be lost. If the country's legal system allows for an adequate response to copyright infringement or other attacks on exam security, losses may be mitigated. For example, countries that are signatories to the Berne convention provide copyright-infringement remedies that are similar to those provided by the United States.

You should add these elements to your list as you consider testing internationally:

5. Indices of fraud, corruption and crime
 - Are commercial and Internet fraud prevalent?
 - Is piracy of exam content rampant?
 - Are bribes common and even expected?
6. Legal protections provided for the protection of intellectual property
 - How quickly are trademarks and copyrights issued?
 - Will copyright registrations of secure tests be honored?
 - Can the rights of copyright holders be enforced?
7. Ability to conduct investigations
 - Do you have resources that can respond to a breach?
 - Can investigators operate legally in the country?
 - Will police and local law enforcement respond to test theft?

It is important to realize that threats and risks vary between countries. Countries where greater corruption, crime and fraud occur will result in higher probabilities of loss to your organization. Risks may be lessened in countries which provide greater ability to investigate and prosecute instances of test fraud.

Solutions and Countermeasures

The best solutions and countermeasures to international testing threats and attacks are found in strengthening exam security processes. In general, the exam security process can be described using a cycle with four critical steps:

- Protection and deterrence
- Detection
- Decision and response
- Review and improvement

Protection and deterrence is facilitated by the analysis of vulnerabilities, threats, attacks and risks that have been described above. Some areas of the world may just be too risky for testing. Other areas may be fine when suitable protective steps are taken.

Detection happens through well-trained proctors, site monitoring, Web monitoring and data forensics analysis. Anonymous tip lines also are valuable for this purpose. All are critical for detecting test security incidents and potential breaches.

Decision and response must be pre-planned and ready to deploy in a timely manner. A security incident response plan is essential. Training is critical for ensuring that your organization is prepared and ready to deal promptly with security breaches. It is imperative to enlist the assistance of local experts who can implement your response within each country.

Review and improvement is critical in order to understand the magnitude of risks and losses. Your organization can improve test security by reviewing test security incidents and determining how to improve protection, detection and response.

Conclusion

When making the decision to expand your program into another country, your organization should consider the size of the opportunity and the impact on test takers:

8. Opportunity and impact of international testing
 - Do you have strategic or economic reasons to begin testing in the country?
 - If entry into the country is delayed, will test takers be excessively penalized?

Whether testing at home or abroad, your program should retain control over the security of the exam administration. You should seriously question the wisdom of testing in areas where this is not possible. The recommended checklist should help you evaluate vulnerabilities, threats, attacks and risks encountered in other countries. This evaluation will guide your selection of specific security measures that help preserve the integrity of exams and generate increased confidence in test administration. After you have taken the proper steps, you will be able to confidently move forward with plans for administering tests internationally.



Dennis Maynes is the chief scientist at Caveon Test Security. He has pioneered several methods for the statistical detection of potential test fraud, including the use of clusters to detect cheat rings and the use of embedded verification tests to detect braindump users. He has conducted more than 450 data forensics projects for more than 50 organizations, including 11 state departments of education, 10 medical programs and 12 information technology certification programs. Maynes holds a master's degree in statistics from Brigham Young University.



Aimée Rhodes is the director of examination security at CFA Institute where she and her team are tasked with maintaining the public confidence in CFA-sponsored designations by ensuring that: 1) exam content is secure; 2) no candidate has an unfair advantage over another; 3) only those individuals truly qualified to pass the exam do so; and 4) the risks inherent in exam programs are identified and addressed appropriately. A lawyer by education, prior to joining CFA Institute Rhodes was the director of test security at ACT.

Putting the Principles of Assessment Engineering into Practice — Impacts for the Design and Analysis of Job Analysis Surveys

By Matthew Schultz and John Mattar

Assessment engineering (AE) is an innovative approach to measurement in which principled test design concepts are used to direct the design and development, as well as the analysis, scoring and reporting of assessment results. The adoption of an AE approach also has implications for the development and analysis of surveys used as part of the practice analysis process. Approaches for analyzing response data and the implications for test blueprinting and design also must be considered from the earliest stages of planning. In this article, we'll examine the work that is considered, starting with the decision to adopt an AE approach as the methodology to drive the practice analysis all the way to the conclusion of the analytic work that starts the finalization of blueprint development. Note that we use "newly licensed" interchangeably with "minimally competent."

Implications for the Development of Practice Analysis Surveys

One implication to survey development is the development of multiple survey statements for relatively finite content areas, created with the goal of identifying both the right level of complexity of the task in practice and also the right level of skill required for the practitioner. This can result in a larger number of statements requiring review by participants than would be seen in more traditional approaches. The need to create and receive ratings on a greater number of task statements can put significant pressure on various stakeholders, particularly content subject matter experts (SMEs) who are charged with developing the task statements. SMEs must be trained in the process of both creating task statements, in some cases delineated

by complexity, as well as providing a foundation in assessment engineering, particularly in the coupling of a work activity (the task) with a skill level (the level of skill required to perform the task). Each task statement is constructed along a proscribed format, as task statements lead to task models, which form the basis for developing test items. Task statements combine a work activity with a verb depicting the skill level that the task is performed at.

In addition to the generation of task statements, there is the need to collect, analyze and summarize survey response data in order to provide comprehensive feedback that informs the test developers and stakeholders of the relative importance of tasks. To satisfactorily handle this greater number of statements, one needs to plan and design a methodology from the earliest stages of the project. Consideration needs to be given to the expected number of statements, the number of statements a participant can reasonably be expected to respond to in the allocated time, the modality for the survey, the required sample sizes required to generate adequate statistical power, the population(s) to be sampled and how to balance statements from various content areas across forms. In addition to the logistical issues, there needs to be a determination regarding not just who is going to provide the ratings, but on what domains participants are going to be asked to provide ratings.

In determining who will participate, there is a need to address, in some cases, multiple perspectives on the tasks performed by the newly licensed. The primary constituencies to sample are frequently the supervisors of the newly licensed and the newly licensed themselves. Each of these populations can provide a different perspective on the work that is done — one may be a better source for information on frequency or to depict the breadth of the work experience, while the other may provide greater clarity regarding relative importance. The relative importance of the various samples considered for inclusion will also drive the resulting sample sizes as well as the analyses that result.

An additional design consideration focuses on what information is most important in determining the test blueprint. Is it more important to understand frequency, or importance/criticality? If there is a question regarding the appropriate skill level required for a given task, various statements across the spectrum can be created in order to ascertain which statement best captures the nature of the task in practice. Each of these can be accomplished, but intentional design is required in order to ensure that sufficient informative data is gathered.

Implications for the Analysis of Surveys

Once survey results are in hand, what does one do with the results? There are many stakeholders interested in the results, but they need assistance in sifting through the numbers and putting them into proper context.

Data cleaning and quality control is the starting point, and that requires a careful and thorough verification of the data. Are the sample sizes sufficient? Are we reasonably sure that the participants responded in a thoughtful manner? Can we be confident that the data quality is such that the conclusions we draw are substantive and generalizable?

What metrics should be employed to summarize the results? This is driven by who is going to work with the data as well as what data has been collected. SMEs need analysis that is informative and readily understandable, making their work easier rather than more complicated. In general, the process of developing a test blueprint requires distilling a large number of statements to a critical set that are most helpful in discerning which candidates should be licensed. For this to happen, many statements (hundreds, perhaps) need to be analyzed and placed within a common context: Which of the statements in question are least reflective of the tasks that are performed? Which are the most?

Considered in tandem with how to analyze the data is the question of what data is available to be included. How are the statements evaluated by survey respondents: on importance, on frequency, on criticality or something else?

A number of options are available and may be considered. For each survey statement, there are one or more summary statistics that can be presented. Do we present them as raw values? Do we combine what may be disparate (or related) statements into an index of some form? Do we scale the data? Do we flag results based on being particularly high, or low, on the domains they're being rated on?

The selected approach to summarizing/flagging data must also consider the populations; if we have supervisors only, we can simply present their data. On the other hand, where we have incumbents as well as supervisors, a determination regarding how to treat these different populations must be decided upon early on. Does one population get a higher weight than the other? What do we do when the different populations provide very different feedback regarding a given statement or test area?

While survey results contribute heavily to the development of test blueprints, they are not the sole drivers, as results are typically presented to panels of SMEs for consideration in the formalization of blueprints. In these situations, how much information is presented to the SMEs providing expert judgment, and how much comparative weight is placed on survey results as contrasted to professional/expert judgment?

Finally, in a large scale, multi-part exam, different teams of people have been utilized in the development of survey statements, survey forms, analysis and summarization, as well as the development of test blueprints and specifications. For activities across these teams, as well as across exam sections, steps should be conducted in as standardized an approach as possible. To that end, the most likely constant in the process, from start to finish, is the participation of psychometrics and test development to provide oversight, planning and methodological guidance throughout the process.

Matthew Schultz is a senior psychometrician at the American Institute of Certified Public Accountants (AICPA), where he currently focuses on psychometric research and test development research. His interests include test design, computer adaptive testing, validity, standard setting, automated scoring and DIF. His work has been presented at annual meetings of the National Council on Measurement in Education (NCME), as well as other regional and national measurement-related meetings. In addition, he has served as a reviewer for NCME.

John Mattar is the director of psychometrics and research at the American Institute of Certified Public Accountants (AICPA), where he oversees all psychometric operations, psychometric research and test development research. He is staff liaison to the AICPA Psychometric Oversight Committee. His interests include test design, validity, standard setting and test development. His work has been presented at annual meetings of the National Council on Measurement in Education (NCME) and the American Educational Research Association (AERA). John has served as a reviewer for annual meetings of NCME, as well as for the journal *Applied Measurement in Education*.

Test Security: A Look Behind the Scenes

By Joy L. Matthews-López, PhD; Ada Woo, PhD; Alan J. Thiemann, JD; Paul E. Jones, PhD; and Jennifer Gallagher

Test security touches nearly every aspect of the testing process. From test design, to score production, to communication with examinees, the overarching umbrella of test security is in play before, during and after tests are administered. The purpose of this article is to highlight security policies and procedures across the test-taking landscape. It also provides a glimpse behind the scenes and shares insights to best practices in the area of test security from both the psychometric and legal perspectives.

Before Test Day

Statisticians warn that analyses cannot fix what has been bungled by poor design. The same warning is often made by psychometricians and attorneys when discussing test security. Common wisdom tells us to prepare for what might be prevented, prepare for what might be detected and prepare for the unexpected. The cornerstone of such preparedness is a defensible test security plan.

A test security plan is a collection of policies and procedures that establish an outline of institutional actions and responses to certain incidents and/or events. Constructed well in advance of a test's administration, a test security plan addresses candidate messaging, non-disclosure agreements and operational issues, such as test design, item types, item exposure and other technical aspects of testing. It requires meticulous documentation of policies and procedures as well as evidence of consistent adherence to them. The worst mistake an organization can make is not having a solid test security plan; the next worse mistake is not applying the plan consistently.

So how can a security breach be prevented? No approach can be 100 percent foolproof, but there are tried-and-true steps that most high-stakes programs know about and take. The most important of these is proactive, intentional, unambiguous, targeted messaging to candidates. Such messaging must be direct and clear regarding roles, rules, expectations and consequences. Beyond the usual messaging to candidates, such as how to register, scope of content and general guidelines, test security messaging addresses the rights and responsibilities of test takers, as well as what procedures, including an appeals process, will be followed in the event of a canceled or invalidated score.

Other proactive steps may include customized non-disclosure agreements for subject matter experts, proctors, internal staff and, of course, examinees. Test development considerations may include development and use of alternative item types less prone to harvesting, copyright of test questions, scoring rubrics and/or answer sheets and formats, use of psychometric models that fit well with high test volumes for quick turnaround of pretest items. It may also be appropriate to adopt operational policies, such as windows-based vs. continuous testing, maximum allowable item exposures, repeater policies, item review flagging policies and procedures, maximum number of attempts, and restrictions on who has access to content as well as secure file transfers and data encryption. Test development staff is in the best position to provide security training to all subject matter experts and to require everyone that comes in contact with content to sign a non-disclosure agreement (NDA). Beyond requiring an NDA with external vendors, test developers may require evidence of security training and awareness from those vendors that handle sensitive content or other intellectual property. Lastly, test developers should proactively establish a validity committee to review incidents and/or flagged records. Doing so strengthens the legal case when scores need to be canceled or invalidated or a score is challenged.

On Test Day

On the day of test administration, numerous security-enhancing steps can be taken; the most obvious of these is proctoring. Regardless of where the test is administered, it should be proctored by a trained test administrator and the examinee's identity verified. Brick and mortar test centers typically offer high-tech options (such as palm vein scanning and digital fingerprinting, video recording of test sessions). Program-approved calculators supplied by test centers are always preferable to examinee-owned calculators, and policies to prohibit use of cellphones or other electronics should be established and enforced. All personal belongings should be left outside the test center or stored in an approved, monitored area of the test center. Personal white boards and markers offer an alternative to scratch paper and examinee-supplied writing instruments, and privacy dividers between computer stations and/or desks provide a barrier to roaming eyes. Most, if not all, test programs require examinees to read and sign non-disclosure agreements both before and upon completion of their test session. In some programs, data from Trojan horse items are used to conduct real-time security analyses. If an examinee's session is flagged as anomalous, then a pre-packaged set of items can be pushed to the test examinee (in real time) so as to minimize the risk of compromising actual test items.

From rigorous check-in procedures, to policies that create a conducive environment for secure testing, to proctoring and surveillance systems, test developers and vendors alike have numerous options to minimize opportunities for examinee misconduct.

After Test Day

After the test is administered, pre-established screening criteria are used to flag anomalous response records, large score differences (for repeat examinees), and unusually low exam scores. In addition, data are typically screened for unusual response time vectors and other markers correlated with misconduct.

Forensics allow records from large groups or cohorts of examinees to be screened for possible collusion, item harvesting, evidence of preknowledge and/or proxy testing. The time needed to run these types of analyses varies from program to program, but it would not be unusual for psychometricians to spend several days running such analyses and interpreting the outcomes. In the event that a test record is flagged, examined and determined to be suspicious, that individual's test may be referred to a validity committee for additional review and possible action. Composition of a validity committee varies by organization, but it usually consists of a senior psychometrician, senior test developer, senior manager, exam security officer and legal counsel.

Ongoing Activities

Finally, there are ongoing activities that can be used to screen for misconduct. Web patrolling, monitoring chat rooms and other social media, and keeping a watchful eye on test prep programs are typical ongoing activities. The program should develop and update letter templates, candidate messaging, websites and NDAs.

Collaborating with peer institutions helps to standardize key policies and procedures. Conducting (and collaborating on) research, and applying research findings to operational policies and procedures produces a body of evidence that can bolster a legal argument (or defense), when necessary.

Legal contributions include but are not limited to the development of cease and desist policies, correspondence templates to communicate with candidates, non-disclosure agreements, and messaging of score cancelation/invalidation notification, as well as participation in formal proceedings.

Lastly, internal (and possible external) audits of an organization's test security infrastructure are periodically conducted to ensure alignment with best practices, industry standards and current legal requirements. Collectively, the combination of good prior planning, solid execution, thorough post-administration screening, and ongoing support from test development and legal staff results in the best infrastructure for test security. Also, it buys, at least to some extent, some peace of mind.

Joy L. Matthews-López, PhD, is a senior psychometrician at the National Association of Boards of Pharmacy (NABP) and an adjunct instructor of mathematics at Ohio University. She received her master's degree in applied mathematics and statistics from NC State and completed her PhD in educational research and evaluation at Ohio University. Her current applications are in the area of licensure testing and her research interests include test security, automated forms assembly, innovative item development and test adaptation.

Paul E. Jones, PhD, is a senior psychometrician at the National Association of Boards of Pharmacy. He received his PhD in psychology at Brigham Young University in 1990. In the 20 years prior to his current position, he served as a psychometric consultant to clients large and small in the information technology and professional credentialing markets providing program and test design, test development, psychometric analysis, standard setting, and auditing services.

Alan J. Thiemann, Esq., has practiced law in Washington, D.C., since graduating from the University of Texas School of Law in 1975. He has been in private practice since 1987 and has spent time as a government lawyer and as in-house counsel to several national trade associations. He is currently the legislative counsel to the Association of Test Publishers and has represented the organization and its members on corporate/nonprofit matters, state and federal legislation/regulations, intellectual property and litigation involving psychological and educational testing.

Ada Woo, PhD, is the director of measurement and testing at the NCSBN. Woo joined the NCSBN psychometric team in June 2008 as a psychometrician, where she was charged with monitoring the psychometric quality of the NCLEX Examinations with regard to reliability, validity and all other aspects related to data integrity. In 2014, she accepted the position of director of measurement and testing, responsible for overseeing all operational aspects of the NCSBN examination programs.

Jennifer Gallagher joined NCSBN in 2005, and in 2008 accepted the position as the operations manager in examinations. In this role, she ensures all NCLEX examinations are administered in a fair, standardized and secure manner through the implementation and development of policies and procedures. She also manages examination operations, security and administration with NCSBN contracted test delivery vendor.

5 Questions to Ask Before Choosing a Remote Proctoring Provider

By Jarrod Morgan

Remote proctoring can open up a world of possibilities for the credentialing community. It brings accessibility and scalability to Internet-based professional testing, making it possible to reach test-takers around the world, and offers a secure alternative to testing centers. Though the online proctoring industry is small, it is rapidly growing, which means those in the professional testing industry have an increasing number of options to choose from. How can you navigate the evolving remote proctoring landscape and choose the provider to best fit your company's needs? Start by asking the following questions:

What type of proctoring does the provider offer?

Remote proctoring companies generally offer one of three approaches to proctoring: automated, record and review, and live proctoring. Fully automated proctoring uses only software to monitor test-takers and determine whether they are cheating. Record and review proctoring offers the option to record the test-taker and then have the recording viewed by trained proctors. By comparison, live proctoring provides a trained professional to monitor the test-taker in real time via Webcam. Each option has its benefits and challenges. Automated options can be implemented at a low cost, but they lack a human factor and may not allow for variations in behavior. Record and review incorporates a human element and usually catches cheating after the exam has been submitted. Live proctoring can deter cheating all together since a live proctor is monitoring the test-taker in real time during the exam.

Can I visit the service facility and learn more about the proctors?

With all of these options, you should also ask who proctors are and how they are trained to proctor test-takers. Regardless of what type of proctoring service you choose, it is also important to have a clear understanding of the proctoring environment. Smaller operations may not be able to provide as much support and scalability as larger service providers. You can learn a lot from taking the time to visit the facility and meet with the teams involved, including quality control, assessment services, programmers and operational managers. Come prepared with a list of additional questions for how the provider can meet your needs.

How quickly can test-takers get help if something goes wrong?

As previously mentioned, each type of proctoring service comes with its own set of challenges. It's essential to have a clear idea of how the service provider will handle challenges when they arise. Whether there is a technical issue or an unexpected concern from the test-taker, it's important to know how the remote proctoring service will provide support. No one wants to wait for a problem to be solved to take an exam. The faster issues get resolved, the fewer complaints you will have to deal with. Ideally, you should look for providers who offer support in real time and answer questions immediately.

What is the provider's track record?

A hiring manager wouldn't be able to hire someone without taking a look at the applicant's previous job experience; the same should be said for deciding on a remote proctoring service provider. Ask for information on their track record, including the number of tests successfully proctored, the number of organizations they work with, and what past clients have to say about their experience, which leads into the last question:

Can I call an organization that works with the provider?

Using the same analogy of the hiring manager, it's a best practice to call an applicant's references to get a better idea of what it is like to work with this individual. It's common for service providers to boast about being the best in their industry, but as the credentialing community knows, it's best to have some sort of measurement for this. Apart from asking about the company's track record, asking for a reference is a great way to get an unbiased view of what it is like to work with the provider.

Remote proctoring offers flexibility for test-takers. For the credentialing community, this allows professional testing to be more efficient and accessible to individuals everywhere. These five questions are a great place to start when deciding on a proctoring service provider, but ultimately, the more information you have on the different options available to you, the easier it will be to find the best solution to fit your needs.



Jarrod Morgan is the co-founder and chief operating officer of ProctorU, which provides secure, live, online proctoring and identity management services in professional testing. ProctorU has 700 partner institutions and has successfully monitored over 1 million exams. Morgan created the initial prototype for the service to fill a need for Andrew Jackson University, where he served as director of technology. He works out of ProctorU's Hoover, Alabama, office where he oversees the operating divisions of the company.

A Psychometrician's Musings on Recertification: What it Means, What You're Really Assessing and How to Ensure Fairness

By Liberty Munson, PhD

Many certification programs require periodic evidence of competence to maintain the use of a credential. On the surface, recertification seems like a simple concept: Candidates need to demonstrate that they continue to possess the knowledge, skills and abilities that are required for successful practice within the content domain. But digging deeper reveals a more complicated story.

It all starts with your definition of recertification. In other words, what is the stated purpose of your recertification program? What are you telling your stakeholders that recertification means? Depending on your definition, recertification likely takes one of the following forms:

- **Continued competence**, means that candidates are engaging in the activities essential to maintain the knowledge, skills and abilities required for practice — that is, they are at least as competent as those earning the credential today. In some cases, this may be more (or possibly less) competent than when they originally earned the credential depending on how the field of practice and the definition of “minimally qualified” has changed during the intervening time period
- **Enhanced competence**, meaning that candidates are demonstrating a higher level of competence than when they originally earned the credential (often based on minimal competence) and potentially higher than those who are entering the field today. A compelling argument can be made for recertification demonstrating enhanced competence in that these candidates have more experience than new entrants and this

would imply a higher level of skill; for many programs, this seems like a reasonable expectation.

Most programs state that the purpose of recertification is to ensure “continued competence” to practice, but the recertification requirements actually necessitate a demonstration of a higher level of skill. For a recertification program to be valid, the requirements must align with its stated purpose or goals. If your recertification requirements require that credential holders demonstrate a higher level of skill than their original certification or than those of new entrants, you’re leveraging recertification as a means for demonstrating “enhanced competence.” This is perfectly acceptable, but the purpose of recertification as “enhanced competence” should be communicated as such to your stakeholders. Because stakeholders need to understand what skills and abilities have been demonstrated by the credential holder, when your recertification process actually assesses increased competence, you should consider renaming the credential or clearly identifying when recertification occurs on transcripts so that stakeholders know that higher levels of competence have been demonstrated. Not only does this ensure the meaning of the credential is clearly communicated, candidates likely benefit from stakeholders knowing that they have demonstrated “enhanced competence.”

If your program truly means “continued competence,” your recertification process should not require the demonstration of a higher level of skills; essentially, candidates need to demonstrate that they continue to be competent at the level required for entry into the field.

This is not to say that they shouldn’t be required to keep their knowledge and skills up-to-date as changes occur (because knowledge of these changes should be expected of new entrants as well), but they should not be expected to demonstrate higher levels of proficiency in the recertification process. One approach to recertification as “continued competence” would be to include a way for them to demonstrate that they still know how to perform at the minimal level to hold the certification and that they are aware of any material changes that have occurred since their last skills assessment.

In the ideal world, the purpose of the recertification program should drive the requirements rather than the requirements ultimately defining the purpose. As you design your recertification program, give careful consideration to what its purpose is and define the requirements accordingly. Some common requirements include: exams, self-assessments, continuing education, peer reviews, submission of work products, etc. How you recertify adds another layer of complexity to the recertification conversation, because little research has been done to demonstrate that methods such as self-assessments, continuing education and peer reviews actually result in practice outcomes that are the basis of recertification (continued or enhanced competence). For example, simply taking a course does not ensure that knowledge and skills transfer to the candidate’s ability to practice or that someone continues to be competent, even if a post-course assessment is required.

Want to add even more complexity to this conversation? Consider recertification programs that give candidates the choice of how to recertify (training, exams, work products, peer reviews, etc.). When you give candidates a choice on how to recertify, you need to ensure that the

evaluation of competence treats all candidates fairly, meaning that the demonstration of competence must be equivalent among candidates regardless of their choice — one choice is not more (or less) difficult than another and that all provide evidence of the same level of competence. For example, is the requirement to pass one or more exams the same as requiring someone to complete a certain number of continuing education credits or provide a portfolio, etc.? Does each option for recertification provide the evidence of the same level of competence? How are you ensuring that you are providing equivalent approaches to evaluating someone's continued or enhanced competence?

These are the questions and considerations that keep psychometricians up at night as they consider the validity, reliability and fairness of recertification programs; what they are intended to do; and how that does or does not align to how the program is actually implemented.

***Liberty Munson** is the principal psychometrician and product quality lead for Microsoft Learning's technical certification program. She is responsible for ensuring that psychometric standards are rigorously applied during all phases of the certification exam lifecycle and that the design and implementation of Microsoft's certification program results in valid and reliable assessments of candidate skills. She received her bachelor's degree in psychology from Iowa State University and her master's and PhD in industrial/organizational psychology with minors in quantitative psychology and human resource management from the University of Illinois at Urbana-Champaign.*