Industrial and Organizational Psychology, **2** (2009), 52–57.
Copyright © 2009 Society for Industrial and Organizational Psychology. 1754-9426/09

UIT or Not UIT? That Is Not the Only Question

DOUGLAS H. REYNOLDS, LAURIE E. WASKO, AND EVAN F. SINAR
Development Dimensions International

PATRICK H. RAYMARK
Clemson University

JOSEPH A. JONES
Development Dimensions International

We believe that "unproctored Internet testing" (UIT) should not be the main focus of concern for workplace psychologists when discussing the implementation and utilization of an Internet-based testing platform. Our basis for this position is that the concept is overly broad and therefore subsequent debate regarding UIT will be imprecise and misleading. Without a more focused set of distinctions among Internet-based testing conditions, useful research and clear guidance for practitioners will be nearly impossible. The term UIT represents a range of test deployment conditions that vary so widely that they preclude accurate and unqualified statements about Internet-based testing. Treating UIT as if it were a consistent set of conditions may polarize opinions and limit progress toward a better understanding of the underlying issues.

Lingering concerns about Internet-based testing cannot be resolved without a broader investigation of the full range of conditions under which candidates access tests over the Internet and a more detailed understanding of the processes by which these conditions may impact important outcomes. In this comment, we build upon the work by Naglieri et al. (2004) and propose a broader framework for understanding test deployment conditions to guide the practice of Internet-based testing.

Correspondence concerning this article should be addressed to Douglas H. Reynolds. E-mail: doug.reynolds@ddiworld.com

Address: Development Dimensions International, 1225 Washington Pike, Bridgeville, PA 15017

Douglas H. Reynolds, Laurie E. Wasko, Evan F. Sinar, and Joseph A. Jones, Development Dimensions International; Patrick H. Raymark, Department of Psychology, Clemson University.

Internet Testing: A Multitude of Questions

Business trends involving globalization, outsourcing of noncore activities, and the drive to reduce administrative costs have helped to fuel the pace of technology adoption across a range of HR practices. These trends have forced selection specialists to adapt to new business practices. The deployment of tests using the Internet is but a subset of the larger trend, and the role of a test proctor is a further subset of the issue.

To be effective as advisors and designers of selection systems, workplace psychologists need to consider a variety of test deployment condition options and must configure these alternatives to meet the organization's business needs while maintaining the integrity of the testing process. Consider the following questions, representative of those that may arise when implementing tests that are distributed and administered via the Internet:

https://doi.org/10.1111/j.1754-9434.2008.01108.x Published online by Cambridge University Press

- What types of test content are most appropriate given the available modes of test administration?
- How do candidate resources and/or individual differences influence their decisions about where to complete the test?
- What strategies can be employed to increase integrity of test content and reduce unintended test exposure?
- What are the abandonment rates at each stage of an Internet-delivered selection process, and how can they be reduced?
- What is the impact of test administration location characteristics on test scores and the candidate experience?

To be sure, the role of a proctor in the testing context is an important consideration in selection system design, but we feel it should be viewed within a broader framework and as only one of numerous factors worthy of discussion and investigation. Proctoring is one selection system design choice to mitigate known and anticipated risks, but by making this the primary variable for debate, we may be missing the forest for this particular tree.

A Broader Framework for Test Deployment Conditions

Simply stated, instead of UIT as a focus area for research as suggested in tables 1 and 2 in Tippins (2009), a broader framework would circumscribe the wide range of deployment conditions that may apply when implementing Internet tests. Such a framework could inform research and guide practice regarding optimal choices for the design of operational selection systems. Figure 1 presents a model designed to outline a range of relevant test deployment conditions. The framework reflects important input, design, moderating, mediating, and outcome factors that are encountered when configuring selection systems. Note that this model, as it is presented, is intended to be neither comprehensive nor prescriptive. The variables described in the model and the potential relationships among them are described in the following sections.

Hiring Context

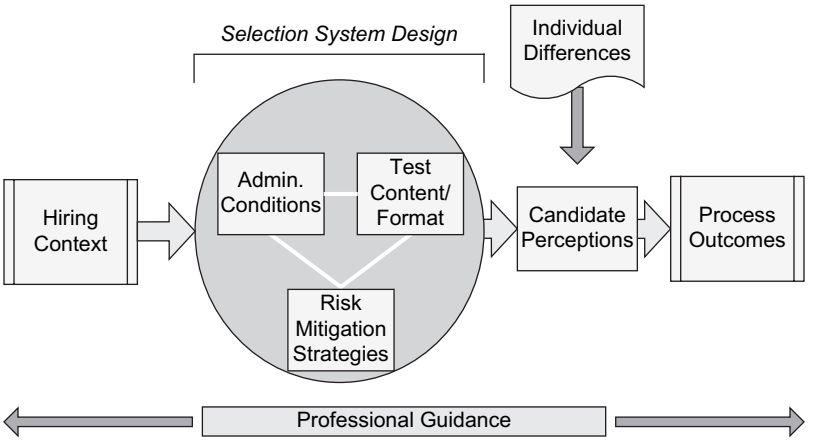
The integration of Internet-based testing into an organization's selection system begins with recognition and explication of the hiring context, examples of which are applicant volume, selection ratio, job level/type, potential positioning of the test within the broader selection system, and other organizational context variables (e.g., Hense, Golden, & Burnett, 2009). These types of contextual variables then lay the foundation for the identification of selection system design alternatives that are more or less feasible for the organization; that is, they help to dictate the ways in which designers can configure a testing platform to best meet the organizational needs and maintain the long-term sustainability of the selection process. For example, if an organization's applicant population includes individuals living in widespread geographic locations, when considering administration condition variables the option of bringing individuals into an on-site testing location may not be practical. Additional hiring context variables varying in their relevance for particular organizations may include the stakes of the testing program for incoming candidates, how the test is to be used (e.g., screen-out vs. compensatory), and the presence of a test orientation program.

Selection System Design

Although the hiring context variables discussed above exist as business realities and as such are relatively stable, choices made by selection system designers should incorporate this context, guide and influence acceptable practice, and optimize the likelihood of achieving the organization's desired process outcomes. We consider three categories of selection system design below.

Test content and format. As noted in several commentaries, the content of a selection battery has implications for the

https://doi.org/10.1111/j.1754-9434.2008.01108.x Published online by Cambridge University Press



Hiring Context

- Stakes for individuals, organizations, and the public
- Test use (e.g., screen-out, hurdle, final-stage selection)
- Job level/type
- Use of test orientation programs
- Applicant volume
- Selection ratio

Test Content & Format

- Cognitive/skills content susceptibility to cheating
- Noncognitive content susceptibility to faking
- Verifiable/nonverifiable content
- Ipsative/normative personality formats
- Long-form/short-form versions
- Open/closed-ended response formats

Individual Differences

- Race/Ethnicity
- Gender
- Age
- Work experience
- Internal vs. External candidate status
- Overqualification status
- Computer experience/ familiarity
- Internet experience/ familiarity
- Experience/familiarity with the tests/selection processes

Process Outcomes

- Test validity
- Changes in test/item variance and difficulty
- Item exposure
- Selection process abandonment rate
- Detected cheating
- Test security breaches
- Recommendation intentions
- Verification test failures
- System return on investment

Administration Conditions

- Presence/absence of a human test proctor
- Presence/absence of others in the test setting
- Convenience of test site or access point
- Access to resources for technical/administrative questions
- Environmental conditions (e.g., workspace, light, distractions)

Risk Mitigation Techniques

- Identification verification methods
- Warnings and honesty declarations
- Data forensics
- Keystroke monitoring
- Verification testing
- Dynamic test presentation (e.g., randomization, CAT)
- Password protection

Candidate Perceptions

- Perceptions of the testing environment
- Information privacy concerns
- Justice perceptions
- Fairness
- Perceptions of user-friendliness/efficiency
- Test anxiety and motivation
- Perceptions of tests and selection processes, in general

Professional Guidance

- APA Standards
- APA Ethical Principles, SIOP Principles
- *ITC International Guidelines on Computer-Based and Internet-Delivered Testing*
- Federal recordkeeping requirements
- Data protection regulations

Figure 1. Proposed model of test deployment conditions.

appropriateness of various administration modes. For example, the choice of cognitive or noncognitive measures may impact the candidate's motivation and ability to cheat or engage in other acts that threaten test security. Because noncognitive content usually lacks clearly correct answers, the use of outside resources during the test process (e.g., collaboration with others, use of external materials) may be perceived to have a lower payoff for cheaters. This consideration is reflected in Arthur, Glaze, Villado, and Taylor (2009) where score inflation strategies of concern for noncognitive tests are no less problematic for proctored when compared to unproctored settings. Test formats (e.g., ipsative or normative personality items) may also

https://doi.org/10.1111/j.1754-9434.2008.01108.x Published online by Cambridge University Press

be chosen in response to particular hiring situations.

Administration conditions. Professional guidelines promote testing under appropriate and adequate conditions (e.g., low noise, few distractions) that are standardized across test takers. When tests are administered over the Internet, some conditions (such as instructions provided to test takers or enforcement of test-level time limits) can be standardized more precisely than with traditional paper-and-pencil tools. Other conditions may vary widely across test takers (e.g., lighting, temperature, Internet speed, the presence of others). Further investigation is merited on the environmental conditions in which applicants are taking tests, and the effect that these conditions have on selection outcomes; in fact, there is some initial evidence that the on-site versus off-site distinction for Internet tests may be more important for some applicant reactions than the proctored versus unproctored distinction (Wasko, Chawla, & Scott, 2007). Further research could better inform the boundaries of fair selection practice when Internet-based tools are used by considering a wider variety of administration conditions, including actual test-taking location, the presence of other test takers, and familiarity with the computer delivering the test. As described by Bartram (2009), the International Test Commission's guidelines delineate four potential administration categories (i.e., open, controlled, managed, supervised mode), and these also provide useful distinctions for further research.

Risk mitigation strategies. Test content security and protection are concerns for any test provider. The use of a proctor is only one approach for addressing these concerns. As evidenced in several articles in the series, additional methods such as monitoring test-taker behavior remotely (e.g., Foster, 2009) or monitoring candidate scores through verification testing (e.g., Burke, 2009) are advancing. Other approaches use various technologies to produce dynamically generated test versions. For example, randomization of items and item responses and the creation of IRT-calibrated parallel forms (e.g., Burke) have been proposed to help protect content. Computerized adaptive testing (CAT) models can also be deployed to reduce testing time and item exposure when tests are broadly distributed.

Candidate perceptions. Each of the aforementioned components of the selection system design may influence candidate perceptions of the test, the testing process, and the hiring organization. Although recent work has begun to examine how administration and environmental conditions may be related to perceptions of the testing environment (e.g., Weiner & Morrison, 2009), much less is known about how candidates react to other design characteristics of Internet-administered tests. For example, the convenience of accessing a test at a location of the candidate's choosing may decrease the amount of test anxiety felt by the candidate and may lead to more favorable impressions of the hiring organization. Another possibility is that the use of password protection may enhance the candidate's perception of test security. One of the reasons why these perception variables require increased attention is that they may provide the causal mechanisms by which the selection system design characteristics influence some of the outcomes associated with the selection process (discussed below).

Individual differences. In addition to the role played by candidate perceptions to mediate the links between selection system design elements and process outcomes, it is important to consider the moderating role of individual difference variables in the relationship between various selection design characteristics and/or candidate perceptions and process outcomes of interest. As an example of such moderating effects, Internet availability varies across levels of socioeconomic status, and these differences in Internet availability and familiarity may impact the candidate's perceived level of comfort and trust toward Internet-based testing processes. Similarly, some administration

https://doi.org/10.1111/j.1754-9434.2008.01108.x Published online by Cambridge University Press

modes may elicit stronger perceptions–outcomes relationships for older candidates as compared to younger candidates. Yet another possibility is that candidates from within the organization may react differently to various aspects of the Internet testing context (e.g., identity controls) than will candidates from outside the organization. In short, we cannot assume that design elements of an Internet-based testing platform will produce similar reactions among all candidate types.

Process outcomes. The impact of various selection system configurations, environmental controls, and mitigation strategies should be considered across a range of outcome variables. Maintaining test validity is one key outcome desired by organizations transitioning to Internet-based administration modes. In addition, "item drift" (i.e., increases in item pass rates over time) has also long been monitored within proctored paper-and-pencil high-stakes testing programs because information about the test will often migrate among test takers; similar measures are critical to monitor for Internet-based tests. Additional outcomes of likely interest for many organizations include selection process abandonment rates, test security breaches, failure rates on verification procedures, candidate intentions to recommend the organization to others, program cost, and overall return on investment for the testing program.

Professional guidance. Although well-informed practice should involve choices made to achieve desired outcomes, these choices should also be guided by ethics, guidelines, and relevant regulations. Pearlman (2009) notes the dominant concerns related to ethics and assessment standards, but Internet deployment also requires further attention to additional rules and regulations. For example, regulations regarding data protection and the transfer and use of data across national borders often apply when the Internet is used to deliver tests (U.S. Department of Commerce, 2008). In addition, federal record keeping rules apply when the Internet is used in personnel selection processes (e.g., U.S. Department of Labor, 2005).

Conclusions

As noted in several commentaries, the Internet testing train has indeed left the station. We are concerned that by labeling and debating the efficacy and ethics of UIT, we will be left in the station as Internet-based HR systems continue down the track without appropriate guidance. Workplace psychologists should perceive an imperative to implement our selection platforms via modern technologies and to understand their strengths, risks, and frailties from an empirically informed foundation. By positioning unproctored Internet testing as something that professionals should be either for or against, the underlying complexities associated with broader test deployment conditions will remain obscured.

References

Arthur, W. A., Glaze, R. M., Villado, A. J., & Taylor, J. E. (2009). Unproctored Internet-based tests of cognitive ability and personality: Magnitude of cheating and response distortion. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2*, 39–45.

Bartram, D. (2009). The International Test Commission guidelines on computer-based and Internet-delivered testing. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2*, 11–13.

Burke, E. (2009). Preserving the integrity of online testing. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2*, 35–38.

Foster, D. (2009). Secure, online, high-stakes testing: Science fiction or business reality? *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2*, 31–34.

Hense, R., Golden, J. H., & Burnett, J. (2009). Making the case for unproctored Internet testing: Do the rewards outweigh the risks? *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2*, 20–23.

Naglieri, J. A., Drasgow, F., Schmit, M., Handler, L., Prifitera, A., Margolis, A., et al. (2004). Psychological testing on the Internet: New problems, old issues. *American Psychologist, 59*, 150–162.

Pearlman, K. (2009). Unproctored Internet testing: Practical, legal, and ethical concerns. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2*, 14–19.

Tippins, N. T. (2009). Internet alternatives to traditional proctored testing: Where are we now? *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2*, 2–10.

https://doi.org/10.1111/j.1754-9434.2008.01108.x Published online by Cambridge University Press

U.S. Department of Commerce (2008). *Safeharbor privacy principles*. Retrieved October 19, 2008, from www.export.gov/safeharbor.

U.S. Department of Labor (2005). Obligation to solicit race and gender data for agency enforcement purposes; final rule. *Federal Register, 70*, 58947–58961.

Wasko, L. E., Chawla, A., & Scott, D. (2007, April). An examination of the opportunities and challenges presented by proctored vs. unproctored testing. In J. A. Weiner (Chair), *The impact of testing conditions on online assessment*. Practice Forum presented at the 22nd Annual Conference of the Society for Industrial and Organizational Psychology, New York.

Weiner, J. A. & Morrison, J. D., Jr. (2009). Unproctored online testing: Environmental conditions and validity. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2*, 27–30.

https://doi.org/10.1111/j.1754-9434.2008.01108.x Published online by Cambridge University Press