Industrial and Organizational Psychology, **2** (2009), 64–68.
Copyright © 2009 Society for Industrial and Organizational Psychology. 1754-9426/09

Moving Beyond the Challenges to Make Unproctored Internet Testing a Reality

ROBERT E. GIBBY
Procter & Gamble

DAN ISPAS
University of South Florida

RODNEY A. MCCLOY
Human Resources Research Organization

ANDREW BIGA
Procter & Gamble

It is clear from Tippins (2009) and the other commentaries that many concerns exist for unproctored Internet testing (UIT), including issues regarding the ethical, legal, and secure use of these tools. It is also clear that many companies and consultancies have moved beyond talking about the issues and have tackled how to make UIT a reality for preemployment selection. The purpose of the present commentary is twofold. First, we will provide a brief overview of how one company, Procter & Gamble (P&G), has considered the issues for UIT in development and validation of its UIT tools. Second, we will discuss the key benefits that compelled P&G to initially incorporate unproctored, Internet-delivered biographical data assessments in 1999 and extend the use of UIT to adaptive cognitive testing in 2008. In the process, we will also be answering several of the open questions from Tippins' Table 2. It is hoped that this discussion will provide additional support and motivate responsible development and use of UIT in industry.

Correspondence concerning this article should be addressed to Robert E. Gibby. E-mail: gibby.re@pg.com
Address: Procter & Gamble, 2 P&G Plaza, TN5-125, Cincinnati, OH 45202
Robert E. Gibby and Andrew Biga, Procter & Gamble; Dan Ispas, Department of Psychology, University of South Florida; Rodney A. McCloy, Human Resources Research Organization (HumRRO).

Overview of P&G Assessment System

To provide a common framework for readers to understand the rest of this commentary, we begin with a brief overview of the standard components of P&G's hurdled selection system. The first step of the P&G process, requirements-based prescreening, involves a set of questions (a) based on job requirements, (b) developed by the hiring manager and recruiter, and (c) delivered online by our candidate management system as part of the candidate's initial job application. Candidates who successfully meet the minimum requirements for the position go on to complete our online, unsupervised biographical data assessment that measures knowledge, skills, and abilities (KSAs) for management, researcher, and office admin job families. This assessment is followed by delivery of our online, adaptive cognitive ability assessment.

Candidates successful through these three online assessments must then complete a proctored paper-and-pencil cognitive ability

https://doi.org/10.1111/j.1754-9434.2008.01110.x Published online by Cambridge University Press

assessment and a series of three structured behavioral interviews conducted by trained and calibrated interviewers. These two supervised components of the selection system provide a deeper look at the candidate's fit with KSAs and allow verification of the online test results.

Managing the Challenges of Using UIT

To manage the security, ethical, and legal challenges described in the other commentaries, we looked to three sources to enable our decision of whether or not—and if so, how—to use UIT tools: our candidates, other companies and consultancies, and the literature.

First, we conducted multiple, extensive candidate reactions research studies to understand acceptance of UIT among thousands of potential and actual candidates. In the first study (Ryan et al., in press), a little over 1,000 students at universities in more than 20 countries were asked to provide reactions to our proposed new testing process, without knowing that it would be used by P&G. The second study (Gibby, Sestak, Pratt, & Irwin, 2006) gained reactions via our career site from several thousand P&G candidates, and the third study (Biga, Pratt, Gibby, & Irwin, 2007) was conducted as part of the research with 150,000+ P&G candidates to calibrate our new cognitive tests.

We began with a focus on candidates because we determined that development of such systems was pointless if candidates would not accept them. The overwhelming response in our research, however, was that they preferred UIT tools to traditional paper-and-pencil methods as long as the former created greater flexibility in taking the tests and faster responses on whether or not they were moving forward. This was a key result because it allowed a level of comfort for there being minimal legal challenge increases coming from candidates after introduction of UIT tools. This level of comfort has held, as we are aware of no legal challenges related to the online use of our assessment tools during the past 9 years.

As a second step, we looked to industry to understand the existing practice for UIT tools. Although there were few benchmarking opportunities in the late 1990s, the prevalence of UIT tools has greatly increased on a global basis, with many companies now employing some form of online screening (e.g., requirements-based prescreening, noncognitive testing, resume mining). Based on our review, we realized that the debate for UIT was not whether it should be used, but rather how it could most effectively be developed, validated, and managed to screen candidates. In discussing the use of UIT by other companies, we were able to develop a list of "best practices" that we were able to refine based on recommendations in the literature.

Given our comfort with candidate acceptance and shared practice of UIT use in industry, we next looked to the literature to develop a best practice list for effective development, validation, and use of UIT that would guide our new assessment system. Consistent with best practices offered in the other commentaries, our list included the following:

- Use of multiple supervised assessments to uniquely verify scores on our online, unproctored biodata, and cognitive assessments.
- Partnership with consultancy and academic leaders in assessment research, development, and practice.
- Extensive and rigorous translation of all assessments (40+ for the cognitive assessments) to enable accurate measurement of the constructs of interest (and not English language ability).
- Concurrent validation of our tests with several thousand employees across the globe under unsupervised conditions for our UIT tools and under supervised conditions for our supervised tests; results for all tests will be at or above industry standards and free of demographic (e.g., age, race, sex) bias.
- Development of cut scores that allow proper flow through of candidates in a way that reduces false negative and false positive results.

https://doi.org/10.1111/j.1754-9434.2008.01110.x Published online by Cambridge University Press

- Development of advanced features for the online, unproctored biographical data assessment, including:
  - Randomized item delivery.
  - Cultural scoring algorithms that enable appropriate measurement of candidates' fit with KSAs.
- Development of advanced features to maximize the precision and security of the online, unproctored cognitive ability assessment, including:
  - Adaptive item delivery of a small subset of the total item pool.
  - Extensive item pools globally developed and screened with more than 150,000 candidates via item response theory calibration and differential item functioning.
  - Item-level, globally calibrated timing to make cheating more difficult.
  - Ongoing live-item research to understand and detect item parameter drift as an indication of item compromise.
  - Ongoing item development and calibration.
  - Item exposure control algorithms.
  - Parallel item pools that can be interchanged, making it difficult for candidates to acquire test content.
- Instruction to candidates that all responses should be made without any help from others.
- Instruction to candidates that their online test results would be verified with additional assessments under supervised conditions.
- Development of an assessment portal system to manage access, consistent messaging, and delivery of multiple UITs to all candidates globally.
- Development of enhanced Web-based delivery systems that minimize the ability of candidates to game the assessments for more time or capture the test questions.

Although the preceding list of best practices is not exhaustive, it is hoped that this list provides insight into how we have tried to manage the security, cheating ethical, and legal risks associated with use of our UIT tools. We challenge the field of industrial–organizational (I–O) psychology to continue its focus on providing answers to industry through research of UIT tools and improved specifications for best practices in using them.

We also hope this list provides a picture of the rigor, resources, and cost required to effectively research, develop, and validate noncognitive and cognitive UIT tools. To reproduce the UIT tools we developed would require a multimillion dollar budget and access to tens of thousands of respondents—both candidates and employees. In addition, extensive I–O psychology and technology resources would be required to specify, design, program, and maintain the online systems. Such an effort would also likely take several years to develop; our system took just over 4 years from initial design to implementation.

Given this overall expense, we realize that our UIT tools are unique in the field and that such a complete system may not be attainable by smaller companies hoping to leverage UIT in their hiring process. We do want to note, however, that components of our system, and especially noncognitive UIT tools, have been successfully developed and used by other companies and are offered by numerous I–O psychology consultancies. Despite such offerings, we decided to make the investment to develop customized online assessments. A common question we receive from internal stakeholders and from industry is, "Why? Why go through all the effort to develop an online process when we could have used a resume-driven, supervised process with less perceived risk?"

Benefits Provided by UIT

Our answer to this question has been to delineate the benefits and rationale for why we continue to feel that incorporating UIT is currently the best available option for screening and selecting candidates. In particular, we focus on the benefits for candidates and the company, and although we present them separately, they are not always mutually exclusive.

https://doi.org/10.1111/j.1754-9434.2008.01110.x Published online by Cambridge University Press

Benefits for Candidates Include: Candidate Benefits

- On-demand, 24/7 access to assessments, offering candidates the ability to manage the testing environment with administration timing that works for them.
- Faster answers about whether the candidates are moving forward in the application process—the primary result asked of us in our candidate' reactions research.
- Standardized delivery of the assessments, absent the issues related to human proctoring of assessments, and ensuring that all candidates receive the same test instruction and preparation.
- Consideration of their application independent of their source (e.g., university campus, career conference), which would not exist with a paper-based system where access to candidates is primarily based on source.
- Providing them the choice of available languages, instead of this choice being managed by the recruiter who may default to English.

Company Benefits

- Ability to manage increasing candidate volumes in a way that consistently considers candidates across geographies and business units—an increasingly important requirement as we have moved from 25,000 applications in 1999 to 500,000+ applications in 2007.
- Ability to cast a wider net for talent, thus permitting us to reach a more diverse pool of candidates.
- Ability to manage a hurdled, multiassessment selection system that provides a more holistic and complete assessment of the candidate's fit with KSAs.
- Ability to deliver messaging to candidates (and consumers) in a controlled way that builds the overall equity of the company.
- Reduced screening times and cost for up-front assessments with real-time score results delivered to recruiters.
- Ability to enrich roles of recruiters and employees who otherwise would have spent time delivering tests.
- Improved cognitive test item security, as described above.
- Ability to consider candidates' cultural backgrounds in scoring noncognitive assessments.

In summary, we believe that our current system—based on available best practices—provides these benefits. We have been able to streamline our global recruiting processes, reducing the number of supervised tests and delivering higher quality candidates to the business in less time. We are also confident that the improved calibration and validity offered by our new tests will lead to improved job and training performance and greater retention of top talent.

Although we have not conducted a formal utility analysis, we have already seen significant cost reductions and improved process efficiencies compared to our prior paper-and-pencil cognitive testing process. As an example, we will deliver approximately 10,000 fewer supervised paper-and-pencil cognitive tests in Japan alone this year. This new process allows us to isolate top talent faster with less cost and fewer resources—both internal and external. An additional key outcome for our UIT tools and systems is that they have been favorably reviewed in audits by government agencies and have stood up to legal challenges. In an era of reduced availability of talent and greater competition for it, we believe this system offers a defensible competitive advantage for P&G. Although our approach is definitely not the answer for every company's hiring process, we are comfortable traveling on the UIT train.

References

Biga, A., Pratt, A. K., Gibby, R. E. & Irwin, J. L. (2007, May). *Cross-cultural differences in applicant reactions to online cognitive ability items.* Symposium

https://doi.org/10.1111/j.1754-9434.2008.01110.x Published online by Cambridge University Press

conducted at the annual meeting of the Society of Industrial Organizational Psychology, New York.
Gibby, R. E., Sestak, N. J., Pratt, A. K., & Irwin, J. L. (2006, May). *Candidate reactions to messaging and assessments of a global selection system*. Practitioner forum conducted at the annual meeting of the Society of Industrial Organizational Psychology, Dallas, TX.
Ryan, A. M., Boyce, A. S., Ghumman, S., Jundt, D., Schmidt, G., & Gibby, R. E. (in press). Going global: Cultural values and perceptions of selection procedures. *Applied Psychology: An International Review*.
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.1754-9434.2008.01110.x Published online by Cambridge University Press