PERSONNEL ASSESSMENT AND DECISIONS

Personnel Assessment and Decisions

Volume 4 | Issue 2 Article 2

12-21-2018

Why Does the Public Sector Resist Unproctored Internet Testing?

Sami Nesnidol
Bowling Green State University, snesnid@bgsu.edu

Scott Highhouse
Bowling Green State University, shighho@bgsu.edu

Why Does the Public Sector Resist Unproctored Internet Testing?

Sami Nesnidol[1] and Scott Highhouse[1]

1. Bowling Green State University

ABSTRACT

Two studies examine public-sector practitioners' concerns about unproctored Internet testing (UIT) for preemployment tests. Study 1 compared public- and private-sector practitioners (n = 66) on possible barriers to UIT adoption (i.e., lack of diffusion, measurement concerns, legal risk, and costs of implementation). Results showed that public-sector practitioners were far less favorably disposed toward implementation of UIT and were more concerned about lack of diffusion, measurement issues, and costs of implementation. Study 2 utilized a policy-capturing design to examine the factors public-sector practitioners consider most important when making simulated decisions about UIT adoption (n = 33). Of the factors examined, test type was found to be the most influential in decisions about implementing UIT.

KEYWORDS

Unproctored Internet testing, public sector, employment testing

Public- and private-sector organizations have different goals and operate under different principles (Aarons, Sommerfeld, & Walrath-Greene, 2009; Rainey, Backoff, & Levine, 1976). Whereas organizations in the private sector have considerable freedom to operate, organizations in the public sector are constrained by rules, traditions, and well-established bureaucratic procedures (Boyne, 2002). In contrast to the private sector, where managers have the ability to hire quickly, it can take several years to create a new position in the public sector—and several months for that position to be filled. Public-sector organizations are also more often required to negotiate with unions when making substantial changes to human resource (HR) practices (Bureau of Labor Statistics, 2017). Also, whereas private-sector hiring is commonly based on perceptions of "fit," public-sector hiring often involves merit-based systems, such as those used in high-stakes police and fire testing.

Given the differences in the hiring environment, it is reasonable to expect that private- and public-sector employers differ in their attitudes about test-administration practices. Especially relevant would seem to be the adoption of new platforms for assessment. With some isolated exceptions (cf., Coffee, Pearce, & Nishimura, 1999), public-sector employers are generally viewed as being slow to adopt new technologies (e.g., Cober, Brown, Blumental, Doverspike, & Levy, 2000). The goal of the present research was to examine sources of public-sector resistance to the adoption of unproctored Internet testing (UIT). First, we compared private-sector assessment experts with public-sector assessment experts to examine (a) whether the perceived difference in adoption rates exists and (b) to compare the two groups on barriers to UIT adoption. In a second study using policy capturing, we examined the relative importance of some of the barriers, as well as the impact of test type, when making simulated decisions about UIT adoption.

Unproctored Internet Testing

UIT involves administering preliminary employment tests online, unsupervised, and outside of a traditional testing environment (Tippins et al., 2006). Benefits of UIT include reduced screening time, reduced costs, standardized delivery, and around-the-clock access to assessments (Gibby, Ispas, McCloy, & Biga, 2009; Tippins, 2009; Tippins et al., 2006). Some have raised concerns, however, about the lack of control over the testing environment, the inability to protect test content or verify applicant identity, and susceptibility of UIT to cheating (Pearlman, 2009; Tippins et al., 2006). Psychometric and score differences between proctored and unproctored online tests, however, appear to be negligible (Arthur, Glaze, Villado, & Taylor, 2010; Le Corff, Gingras, & Busque-Carrier, 2017; Shepherd, Do, & Drasgow, 2003). Additionally, technological advances (e.g., large item pools, browser lockdown, and unauthorized keystroke monitoring) have been developed to help increase test security and limit cheating (Foster, 2009; Gibby et al., 2009). Despite this, it has been our experience that many

Corresponding author:
Sami Nesnidol
Email: snesnid@bgsu.edu

employers still resist using UIT for preemployment testing, especially those in the public sector.

Potential Sources of Resistance

Previous research has suggested several reasons for HR practitioners adopting or resisting a particular selection technique (Harris, Dworkin, & Park, 1990; König, Jöri, & Knüsel, 2011; König, Klehe, Berchtold, & Kleinmann, 2010; Lievens, & De Paepe, 2004). Far less research has been done on factors related to UIT adoption (Lievens & Harris, 2003). From the existing literature, we identified the following factors that may influence UIT adoption.

Potential costs. Employers may be unwilling to use UITs due to the costs of implementation. Previous literature has suggested UIT is likely most beneficial to organizations with large numbers of applicants (Tippins, 2009) where the infrastructure needed to maintain a proctored employment test for all applicants (e.g., computers, proctors, rooms) exceeds that of implementing and maintaining an unproctored test. In contrast, for organizations with fewer applicants, the financial costs associated with UIT (e.g., designing, hosting, securing test information online) may outweigh the benefits. Other factors besides organization size are likely to influence perceptions of costs and benefits. Public-sector employers, for example, may face higher expectations for utility of UIT in order to overcome well-established bureaucratic procedures and union resistance, which are more common to the public sector compared to the private sector (Boyne, 2002; Bureau of Labor Statistics, 2017). In addition, public-sector administrators may feel less decision latitude where costs are concerned, causing them to avoid attempts to move tests to an online format.

Measurement concerns. Test security (e.g., aspects of tests designed to help protect against test content proliferation such as large item pools) and protection against cheating (e.g., limiting access to unapproved helping aids such as search engines or help from another person) are two closely related and popular concerns in the UIT literature (see Beaty, Dawson, Fallaw, & Kantrowitz, 2009; Foster 2009; Gibby et al., 2009; Pearlman, 2009; Tippins, 2009; Tippins et al., 2006). Public-sector practitioners may be particularly concerned with test security and cheating due to fears of item sharing and test sharing in a merit-based selection context (e.g., Sewell & Pringle, 2015).

Legal risks. UIT is a relatively modern technology only achievable with the widespread accessibility of the Internet. Thus, there is little legal precedent regarding the defensibility of UIT for preemployment tests in comparison to older proctored paper-and-pencil and computerized tests. Reluctance to use UIT may stem from fear of litigation, which can result in hefty legal fees and negative publicity for organizations (Terpstra & Honorée, 2016).

Perceived lack of diffusion. Diffusion is the extent to which a test or technique is used by others in the field (Klehe, 2004). Previous research has found diffusion to be a significant factor in determining practitioner willingness to use a particular test (Harris et al., 1990; König et al., 2010). Use of UIT may signal to employers that the procedure is applicable and of sufficient quality to be used in their own organizations. Given the similarity of jobs in different municipalities, public-sector employers may be especially concerned about whether UIT has been used successfully in other localities.

STUDY 1

To examine UIT concerns of private-sector assessment experts versus public-sector assessment experts, we developed a survey assessing sources of resistance to UIT. The survey was based on the potential concerns outlined above (i.e., potential costs, perceived lack of diffusion, legal risks, and measurement issues). Sixty-six HR practitioners (including 45 from the public sector and 21 from the private sector) were recruited through mailing lists and postings on websites directed toward HR practitioners (e.g., HRDIV_NET; LinkedIn), as well as through appeals to regional organizations (e.g., WRIPAC; CIOP) to send the link to their members. The sample was 52% male, 80% Caucasian, with a mean age of 42 years (*SD* = 17). All participants were employed at least part time. The size of participants' organizations varied broadly (< 10 to >10,000) for both public- and private-sector practitioners. Eighty-six percent of private-sector practitioners and 38% of public-sector practitioners reported using UIT in their current place of practice.

Measures

UIT concerns. Concerns about UIT were measured using 18-items designed for this study. Following a description of UIT, respondents were asked to indicate to what extent they agreed with 18 statements. Responses were collected on a 5-point (1 = *strongly disagree*; 5 = *strongly agree*) scale with higher scores indicating greater UIT concerns.

The scale contained four subdimensions corresponding to possible sources of UIT resistance identified in the testing literature: *potential costs* (5 items; α = .82), *perceived diffusion* (4 items; α = .72), *legal risk* (2 items; α = .77), and *measurement issues* (7 items; α = .89). Items are included in Table 1.

Likelihood of use. Likelihood of using UIT was assessed using the item, "Assuming you had total control over the selection process in your organization, how likely would you be to use UIT?" recorded on a 5-point (1 = *very unlikely*; 5 = *very likely*) scale.

RESULTS AND DISCUSSION

An independent-samples t-test showed private-sector practitioners (M = 4.10, SD = 1.14) were significantly more likely to use UIT than public-sector practitioners (M = 3.31, SD = 1.41), $t(47.89) = 2.41$, $p = .020$; $d = .70$. Thus, public-sector practitioners were not only less likely to have experience with UIT, they were also substantially less likely to use UIT if given the opportunity.

Descriptive statistics and correlations between the UIT-concerns subscales are included in Table 2. MANOVA was used to examine whether the concerns varied across occupational setting (public vs. private sector). Results of the MANOVA showed a meaningful difference between public- and private-sector practitioners for the linear composite of the sources, $F(4, 60) = 4.40$, $p = .003$; Wilk's $\Lambda = .77$, partial $\eta 2 = .23$. Specifically, there was a significant difference between public- and private-sector practitioners on potential costs, $F(1, 63) = 18.11$, $p < .001$; partial $\eta 2 = .22$, perceived diffusion, $F(1, 63) = 10.43$, $p = .002$; partial $\eta 2 = .14$, and measurement issues, $F(1, 63) = 4.35$, $p = .041$; partial $\eta 2 = .07$, such that each of these three concerns was greater among public-sector practitioners compared to private-sector practitioners. Perceived legal risk did not differ significantly between the public and private sector, $F(1, 63) = 2.60$, $p = .112$; partial $\eta 2 = .04$. The means and standard deviations for both public- and private-sector practitioners on each of the four factors, as well as the Cohen's d statistics are reported in Table 1. A graphical representation of the means for both groups across the four dimensions is shown in Figure 1.

Thus, the extent to which perceived diffusion, measurement issues, and the potential costs associated with UIT were of greater concern to public-sector practitioners than to private-sector practitioners. We sought to further examine the relative impact of these situational characteristics on public-sector practitioners' decisions about UIT adoption in Study 2.

It should be noted that several study participants mentioned, in an open-ended comments section, that they would be more willing to use UIT for personality rather than cognitive ability tests. We therefore considered this issue explicitly in Study 2—focusing solely on public-sector practitioners.

STUDY 2

Study 1 showed that public-sector practitioners are significantly less likely than private-sector practitioners to use UIT to administer preliminary employment tests. Additionally, Study 1 showed significant differences between public- and private-sector practitioners on measurement issues, perceived diffusion, and potential costs regarding UIT. Building upon Study 1, we used a policy-capturing design to examine the importance of factors identified as particularly concerning to public-sector practitioners in Study 1 (i.e., perceived diffusion and measurement issues stemming from test security and cheating) as well as the effect of test type on decisions to implement UIT.

Potential costs associated with UIT was not examined in Study 2 because the policy-capturing method involves asking people to make simulated decisions based on variation in different aspects of UIT implementation. Thus, the inclusion of program costs in the policy-capturing design would have resulted in scenarios unlikely to occur in the real world (see Aiman-Smith, Scullen, & Barr, 2002). Because of the high cost associated with many test security and cheating prevention methods, it is unlikely that there would be a UIT that possesses both high levels of test security/cheating prevention and is low in cost. It is also notable that this dimension was highly correlated with diffusion in Study 1. Thus, only diffusion was included in Study 2.

We manipulated test type so that the presented scenarios included either general mental ability (GMA), personality, or situational judgment tests (SJTs), which have been successfully implemented in both proctored and unproctored settings (Connell, Arthur, & Doverspike, 2015; Gibby et al., 2009). Both GMA and personality are predictive of performance and commonly measured in preliminary employment tests (Hurtz & Donovan, 2000; Ree, Carretta, & Steindl, 2001; Salgado, 2003; Salgado et al., 2003). Additionally, SJTs can measure a variety of both cognitive abi lities and noncognitive traits (McDaniel, Hartman, Whetzel, & Grubb, 2007), and are commonly used in public-sector testing.

The majority of published literature (e.g., Beaty et al., 2009; Gibby et al., 2009; Hense, Golden & Burnett, 2009; Tippins et al., 2006) agrees with the use of UIT for personality questionnaires and SJTs. Concerns about cheating, however, have resulted in less support for UIT-administered cognitive tests (Pearlman, 2009; Tippins et al., 2006). For these reasons, we expected that public-sector practitioners would be less likely to use GMA tests administered via UIT, compared to SJT and personality tests administered via UIT.

METHOD

Design

This study used a policy-capturing design, which allows researchers to collect information on the relative importance of different factors (e.g., test type) in making a decision. This is done by asking participants for their likely response to multiple scenarios, in which different combinations of the manipulated factors (i.e., cues) are presented. This allows researchers to collect data on multiple judgments quickly, utilizing people involved in making the decisions being examined (Karren & Barringer, 2002). Because policy capturing asks for overall judgments on each

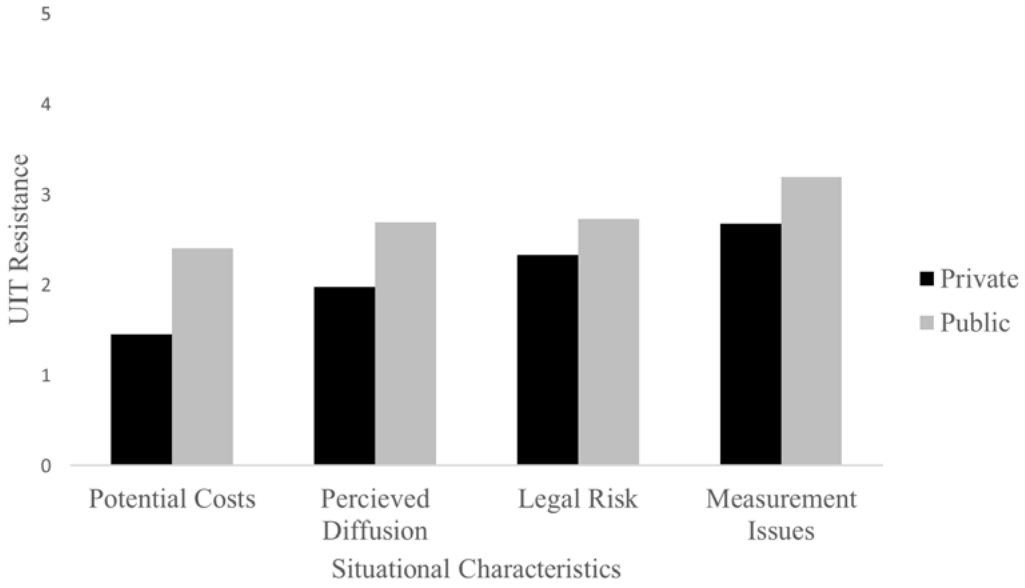
TABLE 1.

UIT Concerns Item and Dimension Means, Standard Deviations, and Cohen's d Values

	Private sector		Public sector		Between groups
Items by subscale	*M*	*SD*	*M*	*SD*	*Cohen's d*
Potential costs	**1.46**	**.45**	**2.28**	**.80**	**1.26**
…does not improve the way we currently screen applicants.	1.90	1.18	2.60	1.25	.58
…would cost too much to implement.	1.38	.50	2.20	1.08	.97
…is beyond our capabilities.	1.33	.73	2.07	.99	.85
…would take too much time to implement.	1.43	.60	2.24	.96	1.01
…would require too much organizational investment.	1.43	.51	2.27	1.10	.98
Perceived diffusion	**1.98**	**.84**	**2.69**	**.81**	**.86**
…is not commonly used.	1.71	.96	2.77	1.03	1.06
…is too novel for my organization.	1.76	1.14	2.42	1.16	.57
…needs to be first used successfully by others.	2.24	1.18	2.78	1.13	.50
…needs to work for others before we would consider it.	2.33	1.28	2.80	1.22	.38
Legal risks	**2.33**	**1.25**	**2.73**	**.77**	**.39**
…lacks legal defensibility.	2.35	1.39	2.69	.97	.28
…lacks legal precedent.	2.33	1.20	2.78	.85	.43
Measurement issues	**2.68**	**.94**	**3.19**	**.89**	**.56**
…does not provide enough advantage over the system we use now.	2.62	1.32	2.44	1.12	-.15
…needs more research on its validity.	2.48	1.37	3.64	.94	.99
…results in bad data on applicants.	2.29	1.10	2.82	1.05	.49
…allows test content to be stolen.	3.48	1.37	3.82	1.05	.28
…results in unreliable scores.	2.35	1.35	2.82	1.11	.38
…data is not as trustworthy.	2.67	1.46	3.22	1.31	.40
…invites cheating.	3.33	1.24	3.56	1.16	.19

Note. Private-sector practitioners $n = 21$. Public-sector practitioners $n = 45$. All items rated on a 5-point (1 = *strongly disagree* to 5 = *strongly agree*) scale.

FIGURE 1.
Comparison of Mean UIT Resistance Between Public- and Private-Sector Practitioners



Results of MANOVA comparing mean UIT resistance scores on each of four situational characteristics across public- and private-sector practitioners. Responses to individual items were recorded on a 5-point 1 = strongly disagree to 5 = strongly agree scale with higher values indicating greater UIT Resistance. Significant differences between public and private sector were found for all situational characteristics except legal risk.

TABLE 2.
Correlations Between UIT-Concerns Subscales

Subscale	*M*	*SD*	*1*	*2*	*3*	*4*
1. Pot. costs	2.03	.80	0.82			
2. Diffusion	2.47	.87	.71**	0.72		
3. Legal risks	2.61	.95	.41**	.29*	0.77	
4. Measurement issues	3.05	.93	.51**	.43**	.71**	0.89

Note. N = 65-66. *p < .05. **p < .01. Cronbach's alpha reported on diagonal.

presented scenario, this results in a more realistic model of real-world decision making compared to participants separately ranking the importance of factors. Compared to traditional self-report methods, policy capturing allows for the collection of more accurate information on the relative importance of the examined factors on decision making by limiting the effects of socially desirable responding (Hitt & Middlemist, 1979; Tomassetti, Dalal, & Kaplan, 2016).

This study used a policy capturing design with three fully crossed decision cues at multiple levels: test type (cognitive ability test, SJT, and personality test), perceived diffusion (high, low), and test security/cheating prevention (high, medium, low). This design resulted in a total of 18 scenarios. Scenarios were presented as short vignettes, with each cue depicted using one sentence or phrase that described the given level of the cue (see Appendix for example scenarios).

Participants and Procedure

Participants were 33 public-sector practitioners currently involved in selection, recruited through a mailing list for human resource and selection professionals (i.e., ipacweb.org). The sample was 39% male, 94% Caucasian (6% Hispanic), with a mean age of 40 (*SD* = 11) years. All participants were currently employed at least 20 hours per week.

Participants were presented with 18 scenarios in ran-

dom order and asked to make simulated decisions about how likely they would be to administer each depicted test via UIT on a 5-point (1 = *very unlikely*; 5 = *very likely*) scale.

RESULTS AND DISCUSSION

Descriptive statistics for the variables of interest can be found in Table 3. A repeated-measures ANOVA was used to examine the main effects of perceived diffusion, test security/cheating prevention, and test type on public-sector practitioners' likelihood of using UIT for preliminary employment tests.

There was a significant main effect of perceived diffusion, $F(1, 30) = 5.67$, $p < .05$, such that practitioners were more willing to use UIT when perceived diffusion was high ($M = 2.40$, $SD = .96$) compared to low ($M = 2.28$, $SD = .93$). There was also a significant main effect of test security/cheating prevention, $F(2, 60) = 20.55$, $p < .05$, such that practitioners were significantly more willing to use UIT when test security/cheating prevention was high ($M = 2.78$, $SD = 1.14$) compared to medium ($M = 2.30$, $SD = 1.00$) or low ($M = 1.95$, $SD = .92$) levels. A main effect of test type was found, $F(2, 60) = 7.87$, $p < .05$, such that practitioners were significantly less likely to use UIT-administered GMA tests ($M = 1.90$, $SD = .79$) compared to SJT ($M = 2.73$, $SD = 1.24$) or personality tests ($M = 2.40$, $SD = 1.25$). No significant interactions were found among the three factors.

For the initial policy-capturing analysis, multiple regression was used to compute the weight (i.e., impact) of perceived diffusion, test security/cheating prevention, and test type on public-sector practitioners' likelihood of using UIT-administered tests. SJT was dropped as a level of test type for this analysis so that test type could be entered into the regressions as one dummy coded variable (0 = *GMA*, 1 = *personality*), therefore resulting in one beta value (i.e., weight) for test type.

The means and standard deviations of the beta values (i.e., weights) of perceived diffusion, test security/cheating prevention, and test type on likelihood of using a UIT-administered test are included in Table 4. The greater the mean value for a factor, the more influential that factor was found to be on practitioners' decisions about UIT-administered tests. Across the sample, test type ($M = .43$, $SD = .31$) was found to have the greatest impact on public-sector practitioners' decisions, compared to test security/cheating

TABLE 3.
Descriptive Statistics of Within-Person Variables

	M	*SD*	Likelihood of using UIT
1. Likelihood of using UIT	2.40	1.37	--
2. Diffusion	.50	.50	.04
3. Test security	1.00	.82	.27*
4. Test type	1.00	.82	.23*

Note. $N = 592 - 594$ *p < .01. Diffusion was dummy coded as 0 = *low* and 1 = *high.* Test security/cheating prevention was coded as 0 = *low,* 1 = *moderate,* 2 = *high.* Test type was coded as 0 = *general mental ability test,* 1 = *personality test,* 2 = *situational judgement test.*

TABLE 4.
Beta Means and Standard Deviations by Cue and Total Variance Explained

	β		*R2*	
	M	*SD*	*M*	*SD*
Diffusion	.12	.20	-	-
TS/CP	.28	.57	-	-
Test type	.43	.31	-	-
All cues	-	-	.72	.18

Note. $N = 26$. TS/CP = *Test security/cheating prevention.* Test type coded as 0 = *GMA* and 1 = *personality.*

prevention (M = .28, SD = .57) and perceived diffusion (M = .12, SD = .20). Of the characteristics examined in this study, therefore, test type is the most influential in public-sector practitioners' decisions about using UIT.

CONCLUSIONS

We found that, compared with public-sector practitioners, private-sector practitioners are more likely to be currently using UIT and are more favorably disposed toward the use of UIT in general. Moreover, public-sector practitioners are more likely than private-sector practitioners to believe that UIT is not worth the investment, is not widely used, and invites cheating. Our policy-capturing study found that public-sector practitioners asked to make simulated implementation decisions favored UIT for personality assessments and SJTs over UIT for cognitive ability tests. Indeed, test type weighed most heavily in decisions to implement UIT, followed by test security/cheating prevention and perceived diffusion. Of the characteristics examined in this study, therefore, test type was the most influential in public-sector practitioners' decisions about using UIT.

These findings support resistance to UIT-administered GMA tests being due in part to concerns over test security and cheating on GMA tests, as suggested by the open-ended comments in Study 1. Our results also show, however, that practitioners weighed test type more heavily than concerns about test security and cheating when making simulated decisions regarding UIT administration. Thus, resistance to UIT-administered GMA tests may also be due to other factors inherent in GMA assessments, such as low face validity for some jobs, which may be seen as likely to provoke negative applicant reactions (Rynes & Connerley, 1993). Additionally, GMA tests are known for possessing a high likelihood of adverse impact, which may result in legal actions being taken against the organization (Viswesvaran & Ones, 2002). Thus, resistance to using UIT-administered GMA tests may be due more to the type of test being given rather than the administration method.

Concerns over test type, test security, and diffusion show the need for public-sector practitioners to stay abreast of the latest empirical research on the validity of assessments administered via UIT, as well as to the need to keep up to date on modern technological advances aimed at enhancing test security. Moreover, testing consultants may need to more effectively communicate the costs and benefits of UIT in the public sector and provide more convincing evidence of viability. Such steps may result in more favorable impressions of UIT for administration of high-stakes tests.

REFERENCES

Aarons, G. A., Sommerfeld, D. H., & Walrath-Greene, C. M. (2009). Evidence-based practice implementation: The impact of public versus private sector organization type on organizational support, provider attitudes, and adoption of evidence-based practice. Implementation Science, 4(1), 83. doi:10.1186/1748-5908-4-83

Aiman-Smith, L., Scullen, S. E., & Barr, S. H. (2002). Conducting studies of decision making in organizational contexts: A tutorial for policy-capturing and other regression-based techniques. Organizational Research Methods, 5(4), 388-414. doi:10.1177/109442802237117

Arthur Jr., W., Glaze, R. M., Villado, A. J., & Taylor, J. E. (2010). The magnitude and extent of cheating and response distortion effects on unproctored Internet-based tests of cognitive ability and personality. International Journal of Selection and Assessment, 18(1), 1-16. doi:10.1111/j.1468-2389.2010.00476

Beaty, J. C., Dawson, C. R., Fallaw, S. S., & Kantrowitz, T. M. (2009). Recovering the scientist–practitioner model: How I-Os should respond to unproctored Internet testing. Industrial and Organizational Psychology: Perspectives on Science and Practice, 2(1), 58-63. doi:10.1111/j.1754-9434.2008.01109

Boyne, G. A. (2002). Public and private management: What's the difference? Journal of Management Studies, 39(1), 97-122. doi:10.1111/1467-6486.00284

Bureau of Labor Statistics. (2017, January 26). Economic new release: Union members summary. Retrieved from https://www.bls.gov/news.release/union2.nr0.htm.

Cober, R. T., Brown, D. J., Blumental, A. J., Doverspike, D., & Levy, P. (2000). The quest for the qualified job surfer: It's time the public sector catches the wave. Public Personnel Management, 29(4), 479-496. doi:10.1177/009102600002900406

Coffee, K., Pearce, J., & Nishimura, R. (1999). State of California: Civil service testing moves into cyberspace. Public Personnel Management, 28(2), 283-300. doi:10.1177/009102609902800209

Connell, M., Arthur, W., & Doverspike, D. (2015, April). Mobile assessment: The horses have left the barn…now what? Workshop presented at the meeting of the Society for Industrial and Organizational Psychology, Philadelphia, PA.

Foster, D. (2009). Secure, online, high-stakes testing: Science fiction or business reality? Industrial and Organizational Psychology: Perspectives on Science and Practice, 2(1), 31-34. doi:10.1111/j.1754-9434.2008.01103

Gibby, R. E., Ispas, D., McCloy, R. A., & Biga, A. (2009). Moving beyond the challenges to make unproctored Internet testing a reality. Industrial and Organizational Psychology: Perspectives on Science and Practice, 2(1), 64-68. doi:10.1111/j.1754-9434.2008.01110

Harris, M. M., Dworkin, J. B., & Park, J. (1990). Preemployment screening procedures: How human resource managers perceive them. Journal of Business and Psychology, 4(3), 279-292. doi:10.1007/BF01125240

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(1), 20-23. doi:10.1111/j.1754-9434.2008.01100

Hitt, M. A., & Middlemist, R. D. (1979). A methodology to develop the criteria and criteria weightings for assessing subunit effectiveness in organizations. Academy of Management Journal, 22(2), 356. doi:10.5465/255595

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. Journal of Applied Psychology, 85(6), 869-879. doi:10.1037/0021-9010.85.6.869

Karren, R. J., & Barringer, M. W. (2002). A review and analysis of the policy-capturing methodology in organizational research: Guidelines for research and practice. Organizational Research Methods, 5(4), 337-361. doi:10.1177/109442802237115

Klehe, U. C. (2004). Choosing how to choose: Institutional pressures affecting the adoption of personnel selection procedures. International Journal of Selection and Assessment, 12(4), 327-342. doi:10.1111/j.0965-075X.2004.00288

König, C. J., Jöri, E., & Knüsel, P. (2011). The amazing diversity of thought: A qualitative study on how HR practitioners perceive selection procedures. Journal of Business and Psychology, 26(4), 437-452. doi:10.1007/s10869-010-9199-9

König, C. J., Klehe, U., Berchtold, M., & Kleinmann, M. (2010). Reasons for being selective when choosing personnel selection procedures. International Journal of Selection and Assessment, 18(1), 17-27. doi:10.1111/j.1468-2389.2010.00485

Le Corff, Y., Gingras, V., & Busque-Carrier, M. (2017). Equivalence of unproctored Internet testing and proctored paper-and-pencil testing of the Big Five. International Journal of Selection and Assessment, 25(2), 154-160. doi:10.1111/ijsa.12168

Lievens, F., & Harris, M. M. (2003). Research on Internet recruiting and testing: Current status and future directions. International Review of Industrial and Organizational Psychology, 18, 131-166.

Lievens, F., & De Paepe, A. (2004). An empirical investigation of interviewer-related factors that discourage the use of high structure interviews. Journal of Organizational Behavior, 25(1), 29-46. doi:10.1002/job.246

McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. (2007). Situational judgment tests, response instructions, and validity: A meta-analysis. Personnel Psychology, 60(1), 63-91. doi:10.1111/j.1744-6570.2007.00065

Pearlman, K. (2009). Unproctored Internet testing: Practical, legal, and ethical concerns. Industrial and Organizational Psychology: Perspectives on Science and Practice, 2(1), 14-19. doi:10.1111/j.1754-9434.2008.01099

Rainey, H. G., Backoff, R. W., & Levine, C. H. (1976). Comparing public and private organizations. Public Administration Review, 36(2), 233-244. doi:10.2307/975145

Ree, M. J., Carretta, T. R., & Steindl, J. R. (2001). Cognitive ability. In N. Anderson, D. Ones, H. K. Sinangil & C. Viswesvaran (Eds.), Handbook of industrial, work, and organizational psychology: Personnel psychology (Vol. 1, pp. 219–232). London/New York: Sage.

Rynes, S. L., & Connerley, M. L. (1993). Applicant reactions to alternative selection procedures. Journal of Business and Psychology, 7(3), 261-277. doi:10.1007/BF01015754

Salgado, J. F. (2003). Predicting job performance using FFM and non-FFM personality measures. Journal of Occupational and Organizational Psychology, 76(3), 323-346. doi:10.1348/096317903769647201

Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F., & Rolland, J. P. (2003). A meta-analytic study of general mental ability validity for different occupations in the European community. Journal of Applied Psychology, 88(6), 1068-1081. doi:10.1037/0021-9010.88.6.1068

Sewell, A. & Pringle, P. (2015, February 21). L.A. County fire officials shared test questions used in hiring, audit finds. Los Angeles Times. Retrieved from http://www.latimes.com/local/california/la-me-fire-cheating-20150222-story.html

Shepherd, W., Do, B. R., & Drasgow, F. (2003, April). Equivalence of proctored versus unproctored online preliminary employment assessments. Paper presented at the 18th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.

Terpstra, D. E., & Honorée, A. L. (2016). Differences in the nature of employment discrimination litigation between private sector organizations and public sector organizations. International Journal of Discrimination and the Law, 16(4), 200-213. doi: 10.1177/1358229116645677

Tippins, N. T. (2009). Where is the unproctored Internet testing train headed now? Industrial and Organizational Psychology: Perspectives on Science and Practice, 2(1), 69-76. doi:10.1111/j.1754-9434.2008.01111

Tippins, N. T., Beaty, J., Drasgow, F., Gibson, W. M., Pearlman, K., Segall, D. O., & Shepherd, W. (2006). Unproctored internet testing in employment settings. Personnel Psychology, 59(1), 189-225. doi:10.1111/j.1744-6570.2006.00909

Tomassetti, A. J., Dalal, R. S., & Kaplan, S. A. (2016). Is policy capturing really more resistant than traditional self-report techniques to socially desirable responding?. Organizational Research Methods, 19(2), 255-285. doi:10.1177/1094428115627497

Viswesvaran, C., & Ones, D. S. (2002). Agreements and disagreements on the role of general mental ability (GMA) in industrial, work, and organizational psychology. Human Performance, 15(1-2), 211-231. doi:10.1080/08959285.2002.9668092

RECEIVED 10/25/17 **ACCEPTED** 10/10/18

Appendix
Example Policy-Capturing Scenarios

Example Scenario 1: Personality test, high diffusion, high test security/cheating prevention

Consider the following description of a selection tool, **administered using unproctored Internet testing**, to answer the question below. Assume acceptable levels of reliability and validity.

- Test type: **Personality**
- Prevalence of use in the public sector: **Common**
- Security features to protect test content and prevent cheating: **Highest level**

How likely are you to use this test administered by unproctored Internet testing?

Example Scenario 2: Cognitive ability test, low diffusion, low test security/cheating prevention

Consider the following description of a selection tool, **administered using unproctored Internet testing**, to answer the question below. Assume acceptable levels of reliability and validity.

- Test type: **General mental ability**
- Prevalence of use in the public sector: **Rare**
- Security features to protect test content and prevent cheating: **Lowest level**

How likely are you to use this test administered by unproctored Internet testing?