

Mobile and Nonmobile Assessment in Organizations: Does Proctoring Make a Difference?

Matthew S. Castillo, Raymond Doe

Department of Psychology, Lamar University, Beaumont, TX, USA
Email: mcastillo21@lamar.edu, rdoe@lamar.edu

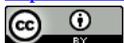
How to cite this paper: Castillo, M. S., & Doe, R. (2017). Mobile and Nonmobile Assessment in Organizations: Does Proctoring Make a Difference? *Psychology*, 8, 878-891.
<https://doi.org/10.4236/psych.2017.86057>

Received: March 7, 2017

Accepted: April 27, 2017

Published: April 30, 2017

Copyright © 2017 by authors and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Advancements in technology have allowed for more efficient methods of testing and assessment. In particular, remotely delivered assessments can be taken on mobile or nonmobile devices in addition to traditional pencil and paper tests. This has led to an increased interest in the comparability of mobile and nonmobile devices on performance outcomes. A variable to consider in performance outcomes on a mobile or nonmobile device is proctoring. There is evidence for both proctored and unproctored conditions leading to better performance outcomes. The present study compared performance on a remotely delivered assessment across mobile and nonmobile devices in proctored and unproctored conditions. Participants were randomly assigned to take a remotely delivered cognitive ability test on either a mobile or nonmobile device in a proctored or unproctored condition. Results indicated that participants tended to perform similarly regardless of the device type or proctoring. Implications are that organizations should consider testing job applicants via mobile devices because performance on a high stakes assessment tends to be similar to testing on a traditional desktop or laptop. Further validation of these results could allow companies to reduce hiring costs by remotely delivering assessments to applicants' own devices.

Keywords

Remote Assessment, Proctoring, Mobile Devices

1. Introduction

Technological advances over the last several decades have led to an increased interest in the remote administration of both simple and complex assessments (Arthur Jr., Doverspike, Muñoz, Taylor, & Carr, 2014). Remotely delivered as-

assessments are those that can be taken online or by using technology as opposed to a written assessment. Today's world no longer confines test taking to pencil and paper format. Tests and assessments can be delivered remotely to both nonmobile and mobile devices including smart phones, tablets, laptops, and desktops. Numerous studies show that the construct validity of internet-based tests is not significantly different than their paper-and-pencil counterparts (Ployhart, Weekley, Holtz, & Kemp, 2003; Potosky & Bobko, 1997; Wilkerson, Nagao, & Martin, 2002). This provides evidence that individuals who take a written assessment should score similarly to those who take an assessment on a mobile or nonmobile device. Why is it important to look at remotely delivered assessments on mobile and nonmobile devices? Recent statistics show that 58% of American adults have a smart phone, 42% own a tablet computer, and 78% own a traditional desktop or laptop (Illingworth, Morelli, Scott, & Boyd, 2015). These numbers have steadily increased over the last few years. A survey of several U.S. companies also showed an increase in mobile and nonmobile device use for pre-employment testing (Fallaw, Kantrowitz, & Dawson, 2012). With the rise in remotely delivered assessments, both individuals and organizations will benefit from research on mobile and nonmobile devices to determine which medium is the best to use for testing. In addition to this information, research concerning the effects of proctoring on performance on high stakes assessments taken on a mobile or nonmobile device will help organizations decide if it is necessary to monitor job applicants on site while they take a test or if assessments can be taken without proctoring.

A mobile device is defined as a hand held, small screen device such as a smart phone or tablet (Jackson, 2013). Mobile devices use Wi-Fi or cellular networking to access the internet. They have an operating system that is not a full-fledged desktop or laptop operating system such as Windows or Linux operating systems. Mobile devices usually have on-screen or attached keyboards for input and typically weigh 2 pounds or less. A nonmobile device is defined as a large screen device such as a traditional desktop or laptop computer. These devices are not portable and usually have a monitor, keyboard, and mouse. A central processing unit (CPU) interprets inputs and executes specified operations. Nonmobile devices contain full-fledged operating systems oftentimes with stronger computing processors than a mobile device (Jackson, 2013). This could lead to performance differences on a remotely delivered assessment depending on the type of device used. Several research studies have looked at the comparability between these two digital mediums in regards to testing and assessment (Sanchez & Branaghan, 2011; Sanchez & Goolsbee, 2010; Schroeders & Wilhelm, 2010). Results indicated that user-interface legibility and user-interface interactivity could influence performance on internet based assessments. The major difference noted between mobile and nonmobile devices is that computers and laptops have large screens, keyboards, and mice or track pads while mobile devices do not. User-interface legibility refers to the ease of reading the material presented on a mobile or nonmobile device. User-interface interactivity refers to the amount of

input a person must do in order to read or understand the material presented on a mobile or nonmobile device, such as scrolling. [Sanchez and Branaghan \(2011\)](#) conducted a study using 34 undergraduate students from a single university. Participants were randomly assigned to read instructions either on a small (mobile) or large (nonmobile) display device. Performance was measured based on an individual's ability to remember the information, or instructions, conveyed on their assigned device type. Results showed that test takers who had to scroll a lot to complete their assessment generally performed worse than those who did not have to scroll a lot. However, rule recall was statistically equivalent regardless of the device type used. Moreover, when participants were allowed to use the mobile device in landscape mode as opposed to portrait mode performance improved. This provides evidence that user-interface interactivity is a larger threat to performance differences across mobile and nonmobile devices compared to user-interface legibility. As a result, it is expected that participants in the mobile device condition will perform better than those in the nonmobile device condition. A limitation of [Sanchez and Branaghan's \(2011\)](#) study is the small sample size used. A larger data set would further validate their findings.

Researchers have suggested that a way to counter the effects of user-interface interactivity on performance across mobile and nonmobile devices is to hold optimization constant by using an assessment that is not optimized for either device ([Huff, 2015](#)). Optimization is important when looking at device usability. An optimized assessment would be one that is designed to maximize the usability for the device that it is taken on. Usability refers to the extent to which a product can be used to achieve goals with effectiveness, efficiency, and satisfaction. For a nonmobile device, usability evaluations would include the computer system, monitor, keyboard, mouse, and any other hardware being used. To address this concern, the present study will use a sample cognitive ability test that was designed by JobTestPrep.com to prepare job applicants to take an actual preemployment test such as the Wonderlic Personnel Test (WPT). JobTestPrep is not affiliated with Wonderlic and the assessment used was not a Wonderlic assessment. Reliability estimates were not provided for the JobTestPrep sample test. Thus, we cannot speak to any psychometric comparability between it and the WPT. Reliability and validity studies on the WPT revealed that it is a test of general intelligence comparable to the Wechsler Adult Intelligence Scale (WAIS) and is based on the Otis Self-Administering Tests of Mental Ability ([McKelvie, 1994](#)). Based on a previous study by [McKelvie \(1989\)](#), the reliability of the WPT was .87. The WPT was developed in 1945 by Eldon Wonderlic and is "one of the most widely used tests of general intelligence" ([Weaver & Bonneau, 1956: p. 127](#)). It is commonly used in personnel selection for pre-employment testing. There are five alternate short forms of the WPT, Forms A, B, D, E, and F. Each consists of 50 items and has a 12-minute time limit. Research on the comparability of these alternate forms revealed that, for the most part, they are psychometrically equivalent. It is noted however that Form B is significantly easier than Form A and Form D is significantly different from Form F ([Kazmier & Browne,](#)

1959). In the present study the sample cognitive ability test accessed online is also a speeded test containing 16 questions with a time limit of 3 minutes 51 seconds. This test can be taken on any device that has access to an internet connection and it was not designed to be taken on a specific device type. The format of the sample test includes answering a question in order to move on to the next question where only one question appears on screen at a time. There are boxes next to each answer choice and a check mark appears in the box when an answer is selected. Screenshots of Question 1 displayed on each device (mobile and nonmobile) are provided (see **Figure 1** & **Figure 2**).

An important variable to consider in testing on mobile and nonmobile devices is proctoring. Much of the research that has been done concerning remotely delivered assessments has focused only on unproctored conditions (Arthur Jr., Glaze, Villado, & Taylor, 2010; Illingworth et al., 2015; Tippins, 2009). The present study expands on this research by including both proctored and not proctored conditions and comparing performance across the two conditions. Proctoring is defined as supervised administration whereas un- or not proctored is defined as unsupervised administration (Weiner & Morrison Jr., 2009). Existing research on the effects of proctoring on performance has yielded conflicting results. For example, Coyne, Warszta, Beadle, and Sheehan (2005) gave a cognitive ability test to 86 students who were randomly assigned to proctored or not proctored conditions and found that participants in the proctored condition scored higher overall. A later study further confirmed the results that performance is higher in proctored over not proctored conditions (Tippins et al., 2006). A limitation of these studies is that both written and remotely delivered assessments were used.

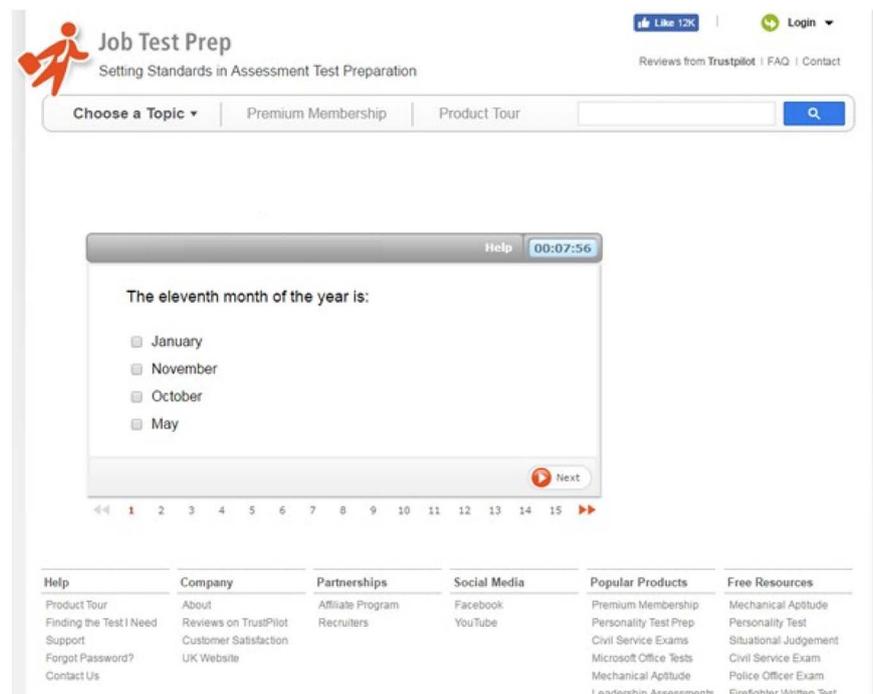


Figure 1. Screenshot of question 1 of the remotely delivered cognitive ability test displayed on a desktop computer (nonmobile device).



Figure 2. Screenshot of question 1 of the remotely delivered cognitive ability test displayed on a smart phone (mobile device).

In contrast, other research has shown that participants in not proctored conditions perform better on cognitive ability tests compared to proctored conditions (Carstairs & Myers, 2009; Murphy & Myers, 2004). Carstairs and Myers (2009) gave an 80 question multiple choice cognitive ability test to 159 undergraduate students and found that participants in the unproctored condition had higher scores than those in the proctored condition. Again both written and remotely delivered assessments were used. The way the test was taken (written versus remotely) could be a confounding variable, so further research is needed. The lack of consistency in research findings on performance differences across proctored and not proctored conditions has made salient the need to further investigate the relationship between proctoring and performance. Results of the present study will add to existing literature by further confirming whether proctoring makes a difference in performance on a remotely delivered cognitive assessment across mobile and nonmobile devices.

High stakes assessments are defined as tests that have major consequences or that are used as the basis for a major decision (Amrein & Berliner, 2002). High stakes are not characteristics of the test itself but consequences of the resulting outcome. The phrase is derived from a gambling term in that “stakes” are a quantity of goods or money that is risked on the outcome of a specific event, such as a hand of poker. Any form of assessment can be used as a high stakes test. In I/O psychology, high stakes assessments commonly take the form of pre-employment tests where an applicant’s performance determines whether he

or she is given a job offer or even short-listed. A moderating factor that is often undiscussed is the speed factor or time allocated for these high stakes assessments. Timed assessments in and of themselves could dissuade a job applicant from searching for the answers to questions, or cheating, due to the fact that they have a limited amount of time to complete the test. A study done by [Arthur Jr. et al. \(2014\)](#) consisted of job applicants who completed a high stakes unproctored remotely delivered assessment for personnel selection purposes. The present study aims to replicate a high stakes remotely delivered cognitive ability assessment in an academic environment. The researcher not only uses an instrument that is commonly used by organizations for making hiring decisions but also emphasizes the need to get all test questions correct. Participants are also informed that the test is timed and their performance would determine if they were selected for a hypothetical job conducting research in the Department of Psychology. In the absence of a valid job offer, this design approximates the conditions of a high stakes assessment within the limit of ethical standards. As a result, it is expected that participants will perceive the situation as high stakes based on the information provided to them in the consent form as well as the verbal instructions given by the researcher.

The use of mobile or nonmobile devices for assessment purposes may result in an underrepresentation of people in lower socioeconomic statuses or worse performance for those who are unfamiliar with technology ([Arthur Jr. et al., 2014](#); [Pearlman, 2009](#)). The present study overcomes this limitation by using a diverse college student population that has at least some familiarity with technology. Furthermore, we aim to expand on the original research done by [Arthur Jr. and his colleagues \(2014\)](#) concerning the use of mobile devices in high stakes remotely delivered assessments. Specifically, we introduce the variable of proctoring to determine whether there are performance differences on a cognitive ability test across mobile and nonmobile devices when being monitored versus not being monitored. The inconsistency of results in the extant literature created a need to explore if performance tends to be higher for a particular device type or proctor condition.

2. Method

2.1. Participants

Undergraduate students taking lower-level psychology courses were used for this study after gaining institutional review board (IRB) approval #73416149. Participants volunteered to take part in the study by signing up via the online SONA system in exchange for research participation credit. A total of 100 participants ($n = 100$) took part in the study. The sample was 84% women ($n = 84$) and 16% men ($n = 16$). Ages ranged from 18 to 67 where 19 was the most frequently occurring age and the average age was 22.

2.2. Materials

A sample pre-employment test was used as the high stakes cognitive ability as-

assessment delivered remotely to mobile and nonmobile devices. The test was accessed via [JobTestPrep \(2017\)](#). The questions and question order remained the same for each participant. A desktop running Windows was used as the nonmobile device. A researcher owned touch screen android phone was used as the mobile device.

2.3. Procedure

After gaining IRB approval, participants signed up for the study using the online SONA system. They were randomly assigned to either the mobile or nonmobile device condition as well as the proctored or not proctored condition prior to arriving to the lab at their designated time. All participants received instructions informing them that they were about to take a cognitive ability assessment. They were also told to try their best to get all questions correct because their performance would determine if they were short-listed for a job in the Department of Psychology. The instructions indicated whether a proctor would be present in the room with the participant or not and what device they would be using. Once the participant was read the instructions by researchers he or she went to the link provided above, navigated to the test instructions page, and began the timed assessment. In the proctored conditions, a proctor sat in the room with the participant. In the unproctored conditions, participants were taken to an empty research room and left alone to take the assessment. For the unproctored condition, participants were told to come get the researcher from a room down the hall when the assessment was complete. When the test was completed a score report appeared on screen and the researcher showed the participant how they performed as well as the questions they missed while recording this information on a coding sheet. Finally, participants were debriefed by being informed that there wasn't an actual job to be offered, their performance on the test had no relationship to their intelligence, and that the purpose of the experiment was to compare differences in test performance on mobile and nonmobile devices in proctored and not proctored conditions. Information concerning the on-campus psychology clinic was also provided in case participants were upset about their experience during the study. This concluded the experiment and participants were shown out of the research lab.

3. Results

Performance differences across mobile and nonmobile devices in proctored and unproctored conditions were evaluated using a Pure Between Factorial Analysis of Variance (ANOVA). The 2×2 design included two between-subjects' independent variables (IV's). The first IV was device type where participants were randomly assigned to take the cognitive ability assessment on either a mobile or nonmobile device. The second IV was proctoring where participants were randomly assigned to either have a proctor in the room with them while they took the assessment or not. Finally, the dependent variable was performance on the cognitive ability assessment as indicated by the normalized score given once the

test was completed. Prior to the ANOVA, the assumption of homogeneity of variance was verified using Levene's Test of Equality of Error Variances, $F(3, 96) = .512, p = .675$. The failure to reject the null hypothesis indicated that group variances were, in fact, homogenous. A completely randomized factorial ANOVA did not reveal a significant main effect of device type, $F(1, 96) = .126, p = .723, \eta^2 = .0002$. Participants in the mobile device group ($M = 3.48, SD = 1.73$) tended to perform similarly to those in the nonmobile device group ($M = 3.36, SD = 1.63$). The ANOVA also did not reveal a significant main effect of proctoring, $F(1, 96) = .350, p = .555, \eta^2 = .0007$. Participants in the proctored condition ($M = 3.52, SD = 1.79$) tended to perform similarly to those in the unproctored condition ($M = 3.32, SD = 1.56$). Finally, there was no significant interaction effect of device type x proctoring, $F(1, 96) = .350, p = .555, \eta^2 = .0007$ (see **Table 1** for a summary of the results).

Participants tended to perform similarly regardless of the device type they were assigned to and whether a proctor was in the room with them when they took the cognitive ability assessment or not (see **Figure 3** for a graph of the cell means). Posteriori effect sizes were calculated for all effects using Eta squared for device type, proctoring, and the interaction of device type and proctoring. According to **Levine and Hullett (2002)** Eta squared values that are less than .01 should be interpreted as a small effect size.

Table 1. ANOVA summary table testing the effects of device type and proctoring on performance score.

Source	SS	df	MS	F
Device Type	.360	1	.360	.126
Proctoring	1.000	1	1.000	.350
Device Type × Proctoring	1.000	1	1.000	.350
Error	274.000	96	2.854	
Total	276.360	99		

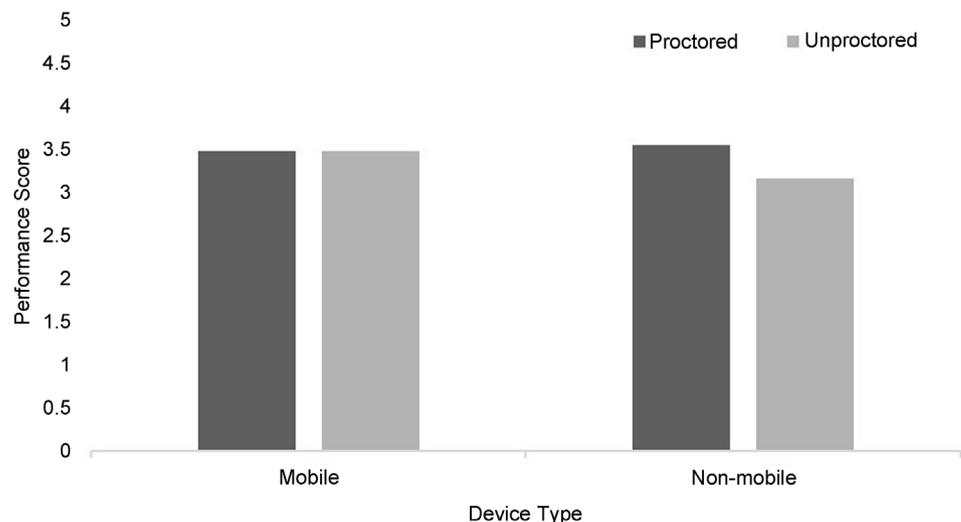


Figure 3. Performance score as a function of device type and proctoring.

4. Discussion

The purpose of this study was to test whether proctoring significantly impacted performance scores on a remotely delivered assessment to mobile and nonmobile devices. Results indicated that performance on a remotely delivered cognitive ability test tended to be similar across mobile and nonmobile devices as well as proctored and unproctored conditions. Taking a high stakes remotely delivered assessment on a mobile device tended to yield similar scores to taking the assessment on a nonmobile device. Taking a high stakes remotely delivered assessment in a proctored condition also tended to yield similar results to taking the assessment in an unproctored condition. It is important to note that these results do not insinuate causation and care should be used in interpreting the findings. For example, there are possible systematic errors underlying the current findings due to the difficulty of simulating high stakes among non job applicants in a controlled environment. Eta-squared values indicated small effect sizes suggesting a low practical significance. Although the intent was to simulate two ontologically distinct conditions, it is fair to argue that the control condition may still have felt some elements of the experimental condition. Participants may not have believed that the test was completely not proctored in the unproctored lab condition. The concept of proctoring can be subjective and varies from context to context. For example, test takers may have a feeling of being watched indirectly when taking a remotely delivered assessment because their information could be traced over the Internet. However, measures were taken to best simulate an unproctored condition in an organizational setting since there is no way to guarantee if someone other than the prospective job applicant is taking the assessment in an unproctored field setting. The present study focused specifically on proctored and unproctored lab conditions which are comparable to an organizational setting.

Organizations can use the above information when making considerations for pre-employment testing. Tests and assessments can be delivered remotely to devices such as smart phones, tablets, laptops, and desktops. Moreover, a survey of several U.S. companies shows an increase in the use of mobile and nonmobile device for pre-employment testing (Fallaw, Kantrowitz, & Dawson, 2012). Companies can choose to test applicants on their mobile devices which could increase the applicant pool seeing as how prospective employees would not be required to access a traditional desktop or laptop computer to take an assessment. As noted earlier at least 58% of American adults own a smart phone and this percentage could be larger among the current U.S. labor force (Illingworth et al., 2015). It may be easier for an applicant to take a pre-employment test on their cell phone rather than on a desktop or laptop. This, in turn, could attract more applicants to apply for jobs which lead to more job offers and positive outcomes for organizations such as decreased recruiting costs.

Decisions can also be made regarding whether to proctor applicants while they take a remotely delivered assessment or not. Much of the research that has been done concerning remotely delivered assessments focuses only on unproc-

tored conditions (Arthur Jr. et al., 2010; Illingworth et al., 2015; Tippins, 2009). The present study did not find statistically significant performance differences on a remotely delivered assessment across a proctored and unproctored lab condition. Previous research yields conflicting results regarding the effect of proctoring on assessment performance. There is evidence that proctored and unproctored conditions tend to have better performance over the other (Carstairs & Myors, 2009; Coyne et al., 2005; Murphy & Myors, 2004; Tippins et al., 2006). However, we were unable to replicate these findings. The present study indicated that participants perform similarly whether a proctor is present in the room with them when they take an assessment or not. Allowing applicants to take pre-employment tests in unproctored environments could save companies a considerable amount of time and money in the hiring process.

The devices used in this study could have affected the results seeing as how participants may have been unfamiliar with the researcher provided Dell desktop or android device. Allowing a familiarity period with the randomly assigned device before starting the assessment may yield different results. It would be interesting to see if an iPhone or other specific type of mobile device would show an increased variability in performance scores compared with the Motorola android device used in this study. Conditions where participants take an assessment on their own device rather than one provided by researchers may also yield different results. Previous research has indicated that user-interface legibility and user-interface interactivity could influence performance on internet based assessments (Sanchez & Branaghan, 2011; Sanchez & Goolsbee, 2010; Schroeders & Wilhelm, 2010). Results showed that test takers who had to scroll a lot on a mobile or nonmobile device to complete their assessment generally performed worse than those who did not have to scroll a lot. One way to counter the effects of user-interface legibility and interactivity on performance scores across mobile and nonmobile devices is to hold optimization constant by using an assessment that is not optimized for either device (Huff, 2015). The present study used a remotely delivered cognitive ability test that was not optimized to be taken on a particular device. Moreover, participants did not have to scroll during the assessment to read a question and its answer choices. Therefore, even though participants were not allowed a familiarity period with their assigned device type before taking the assessment it is possible that the format of the remotely delivered test negated the above concerns.

The assessment used for this experiment had 16 questions and was timed for around four minutes. Many participants did not complete the assessment before time ran out. It is possible that the time factor impacted results and caused a lack of variation in performance scores. Questions that were not attempted were still counted wrong which negatively affected a participant's normalized score. Increasing the time limit or comparing scores on the online version of the test used in this study with an untimed cognitive ability assessment may yield different results. Previous research concerning speeded cognitive ability tests has focused primarily on retest effects across two test administrations (Arthur Jr. et al., 2010;

Nye, Do, Drasgow, & Fine, 2008). This research design was not used in the present study; however, the time factor was necessary to simulate a high stakes assessment. Moreover, speeded ability tests might be one way alleviate malfeasance concerns with remotely delivered assessments.

Arthur Jr. and his colleagues (2010) noted that the absence of proctors could create a permissive environment for cheating. The unproctored conditions used in this study were simulated in a lab setting. Although researchers left participants alone in a room to take the cognitive ability assessment who were randomly assigned to an unproctored condition, they were still ostensibly nearby. As a result, participants may not have truly felt like they were not being monitored. It is not clear whether the notion of being watched or monitored can be completely ruled out in unproctored lab settings. The ideal unproctored environment for our study would be to remotely deliver the cognitive ability assessment to participants' own mobile or nonmobile devices so that they can take the test outside of the lab without being proctored. However, adopting this field approach also raises questions of who is actually taking the assessment. While the current study did not include conditions outside of the lab setting subsequent studies will. Moreover, we assert that the nature of the unproctored condition should not affect overall performance scores due to the fact that the test used in this study was timed. As noted previously, timed tests may dissuade participants from cheating because they do not have adequate time to do so. A future study will include both timed and untimed assessments to empirically support the notion that malfeasance is more likely with untimed tests. We will be able to isolate malfeasance by comparing performance scores on a timed assessment to scores on an untimed assessment in lab and outside of lab settings.

Finally, the simulated high stakes setting in this study may not have been perceived by participants the way we intended them to. We informed participants that the assessment they were about to take was commonly used by many organizations to select candidates for entry level jobs. We repeatedly stressed the importance of participants doing their best on the assessment and attempting to answer all questions correctly. Finally, we told participants that their performance on the assessment would determine if they were short-listed for a hypothetical job. There was no tangible reward to give to participants who performed well on the assessment so they may not have taken the test as seriously as they would in a real pre-employment test situation. Simulating high stakes for jobs that individuals have not solicited directly is a challenge and highly subjective. The use of an incentive or reward should approximate a high stakes setting or at best increase motivation to perform well on an assessment. In a similar study, Mueller-Hanson, Heggstad, and Thornton (2003) compared a control group to an incentive group where participants were told that their performance on a personality assessment would determine if they qualified to enter a drawing to receive \$20. Results indicated that those in the incentive group tended to perform better than those in the control group who were not enticed by a reward. Although we simulated a high stakes assessment as closely as possible a subse-

quent study will include some type of reward or incentive to ensure that participants take the test seriously and truly perceive the situation as high stakes. Future studies should also consider comparing results on a high stakes assessment across individuals who are offered a hypothetical job and individuals who apply for an actual job. Overall, the present study did not find performance differences on a high stakes, remotely delivered assessment to mobile and nonmobile devices in a proctored and unproctored lab setting. Despite a few limitations, this study has provided further support for the notion that proctoring may not significantly affect performance scores which will help organizations determine the utility of unproctored Internet testing in hiring processes and improve our knowledge of remotely delivered assessments.

Acknowledgements

The authors would like to thank Dr. Edythe Kirk for her useful comments and suggestions.

References

- Amrein, A. L., & Berliner, D. C. (2002). High-Stakes Testing, Uncertainty, and Student Learning. *Education Policy Analysis Archives*, *10*, 1-74.
<https://doi.org/10.14507/epaa.v10n18.2002>
- Arthur Jr., W., Doverspike, D., Muñoz, G. J., Taylor, J. E., & Carr, A. E. (2014). The Use of Mobile Devices in High-Stakes Remotely Delivered Assessments and Testing. *International Journal of Selection and Assessment*, *22*, 113-123.
<https://doi.org/10.1111/ijsa.12062>
- 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*, 1-16. <https://doi.org/10.1111/j.1468-2389.2010.00476.x>
- Carstairs, J., & Myers, B. (2009). Internet Testing: A Natural Experiment Reveals Test Score Inflation on a High-Stakes, Unproctored Cognitive Test. *Computers in Human Behavior*, *25*, 738-742.
- Coyne, I., Warszta, T., Beadle, S., & Sheehan, N. (2005). The Impact of Mode of Administration on the Equivalence of a Test Battery: A Quasi-Experimental Design. *International Journal of Selection and Assessment*, *13*, 220-224.
<https://doi.org/10.1111/j.1468-2389.2005.00318.x>
- Fallow, S. S., Kantrowitz, T. M., & Dawson, C. R. (2012). *2012 Global Assessment Trends Report*. http://www.Shl.com/assets/GATR_2012_US.pdf
- Huff, K. C. (2015). The Comparison of Mobile Devices to Computers for Web-Based Assessments. *Computers in Human Behavior*, *49*, 208-212.
- Illingworth, A. J., Morelli, N. A., Scott, J. C., & Boyd, S. L. (2015). Internet-Based, Unproctored Assessments on Mobile and Non-Mobile Devices: Usage, Measurement Equivalence, and Outcomes. *Journal of Business Psychology*, *30*, 325-343.
<https://doi.org/10.1007/s10869-014-9363-8>
- Jackson, W. (2013). *Just What Does NIST Consider a Mobile Device?* National Institute of Standards and Technology Special Publication, 800, 124.
https://gcn.com/articles/2013/06/27/nist-mobile-device-definition.aspx?admgarea=TC_Mobile

- JobTestPrep (2017). *Wonderlic Sample Practice Test*. TestPrep-Online. <https://www.jobtestprep.com/wonderlic-sample-test>
- Kazmier, L. J., & Browne, C. G. (1959). Comparability of Wonderlic Test Forms in Industrial Testing. *Journal of Applied Psychology*, *43*, 129-132. <https://doi.org/10.1037/h0045688>
- Levine, T. R., & Hullett, C. R. (2002). Eta Squared, Partial Eta Squared, and Misreporting of Effect Size in Communication Research. *Human Communication Research*, *28*, 612-625. <https://doi.org/10.1111/j.1468-2958.2002.tb00828.x>
- McKelvie, S. J. (1989). The Wonderlic Personnel Test: Reliability and Validity in an Academic Setting. *Psychological Reports*, *65*, 161-162. <https://doi.org/10.2466/pr0.1989.65.1.161>
- McKelvie, S. J. (1994). Validity and Reliability for an Experimental Short Form of the Wonderlic Personnel Test in an Academic Setting. *Psychological Reports*, *75*, 907-910. <https://doi.org/10.2466/pr0.1994.75.2.907>
- Mueller-Hanson, R., Heggstad, E. D., & Thornton, G. C. (2003). Faking and Selection: Considering the Use of Personality from Select-In and Select-Out Perspectives. *Journal of Applied Psychology*, *88*, 348-355. <https://doi.org/10.1037/0021-9010.88.2.348>
- Murphy, K. R., & Myers, B. (2004). *Statistical Power Analysis: A Simple and General Model for Traditional and Modern Hypothesis Tests* (2nd ed.). Mahwah, NJ: Erlbaum.
- Nye, C. D., Do, B., Drasgow, F., & Fine, S. (2008). Two-Step Testing in Employee Selection: Is Score Inflation a Problem? *International Journal of Selection and Assessment*, *16*, 112-120. <https://doi.org/10.1111/j.1468-2389.2008.00416.x>
- Pearlman, K. (2009). Unproctored Internet Testing: Practical, Legal, and Ethical Concerns. *Industrial and Organizational Psychology*, *2*, 14-19. <https://doi.org/10.1111/j.1754-9434.2008.01099.x>
- Ployhart, R. E., Weekley, J. A., Holtz, B. C., & Kemp, C. (2003). Web-Based and Paper-and-Pencil Testing of Applicants in a Proctored Setting: Are Personality, Biodata, and Situational Judgment Tests Comparable? *Personnel Psychology*, *56*, 733-752. <https://doi.org/10.1111/j.1744-6570.2003.tb00757.x>
- Potosky, D., & Bobko, P. (1997). Computer versus Paper-and-Pencil Administration Mode and Response Distortion in Noncognitive Selection Tests. *Journal of Applied Psychology*, *82*, 293-299. <https://doi.org/10.1037/0021-9010.82.2.293>
- Sanchez, C. A., & Branaghan, R. J. (2011). Turning to Learn: Screen Orientation and Reasoning with Small Devices. *Computers in Human Behavior*, *27*, 793-797.
- Sanchez, C. A., & Goolsbee, J. Z. (2010). Character Size and Reading to Remember from Small Displays. *Computers and Education*, *55*, 1056-1062.
- Schroeders, U., & Wilhelm, O. (2010). Testing Reasoning Ability with Handheld Computers, Notebooks, and Paper and Pencil. *European Journal of Psychological Assessment*, *26*, 284-292. <https://doi.org/10.1027/1015-5759/a000038>
- Tippins, N. T. (2009). Internet Alternatives to Traditional Proctored Testing: Where Are We Now? *Industrial and Organizational Psychology*, *2*, 2-10. <https://doi.org/10.1111/j.1754-9434.2008.01097.x>
- Tippins, N. T., Beaty, J., Drasgow, F., Gibson, W. M., Pearlman, K., & Segall, D. O. (2006). Unproctored Internet Testing in Employment Settings. *Personnel Psychology*, *59*, 189-225. <https://doi.org/10.1111/j.1744-6570.2006.00909.x>
- Weaver, H. B., & Bonneau, C. A. (1956). Equivalence of Forms of the Wonderlic Personnel Test: A Study of Reliability and Interchangeability. *Journal of Applied Psychology*, *40*, 127-129. <https://doi.org/10.1037/h0047065>

Weiner, J. A., & Morrison Jr., J. D. (2009). Unproctored Online Testing: Environmental Conditions and Validity. *Industrial and Organizational Psychology*, 2, 27-30.
<https://doi.org/10.1111/j.1754-9434.2008.01102.x>

Wilkerson, J. M., Nagao, D. H., & Martin, C. L. (2002). Socially Desirable Responding in Computerized Questionnaires: When Questionnaire Purpose Matters More Than the Mode. *Journal of Applied Social Psychology*, 32, 544-559.
<https://doi.org/10.1111/j.1559-1816.2002.tb00229.x>



Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact psych@scirp.org