Intelligence 67 (2018) 67–75

Contents lists available at ScienceDirect

Intelligence

journal homepage: www.elsevier.com/locate/intell

Unproctored internet-based device-type effects on test scores: The role of working memory

Winfred Arthur Jr*, Nathanael L. Keiser, Ellen Hagen, Zach Traylor

Texas A&M University, United States

ARTICLE INFO

Keywords:
Unproctored internet-based testing
Mobile device
Nonmobile device
Working memory
Cognitive ability
Noncognitive constructs

ABSTRACT

Despite the burgeoning use of unproctored Internet-based testing (UIT) in employment-related assessments, there have been limited theoretical and empirical advancements to explain the effects of technology on test and assessment outcomes. Indeed, this issue is potentially germane to all researchers and practitioners who use UIT assessments in their research and practice. To address this gap, Arthur, Keiser, and Doverspike (2017) advanced the Structural Characteristics/Information Processing (SCIP) framework as a means for psychologically conceptualizing the effect of UIT device-types on test scores. The present study examined the working memory (WM) propositions advanced by the SCIP framework. Participants were randomly assigned to a desktop computer or smartphone condition to complete a general mental ability (GMA), and personality (i.e., agreeableness, conscientiousness) measure on either a desktop computer (n = 174) or smartphone (n = 173). All participants also completed a WM measure on a desktop computer. The results provide initial support for some of the SCIP framework propositions in that as hypothesized, the WM/GMA, and WM/completion time relationships were stronger for assessments completed on smartphones, compared to desktop computers; in contrast, the WM/personality relationships were weak and did not generally differ across device types. Consequently, this study offers an initial empirical test of the SCIP framework; however, further research is needed to examine additional aspects of the framework, including the role of the other information processing variables (i.e., perceptual speed and visual acuity, psychomotor ability, and selective attention) advanced by the framework.

1. Introduction

Technology continues to play an increasingly important role in industrial-organizational (I-O) psychology and human resource management (HRM), particularly in the context of personnel selection and assessment (Stone, Deadrick, Lukaszewski, & Johnson, 2015). However, it would seem that the rapid rate of technological change has resulted in practice outpacing science (Arthur, Doverspike, Kinney, & O'Connell, 2017; Morelli, Potosky, Arthur, & Tippins, 2017). This is particularly salient in the realm of unproctored Internet-based testing (UIT) where the associated advantages, specifically, the ability of test-takers to test anywhere and at anytime, have resulted in a burgeoning increase in the use of mobile devices to complete these assessments (Arthur, Doverspike, Muñoz, Taylor, & Carr, 2014; McClure Johnson & Boyce, 2015) because UIT also by definition, in principle, gives test-takers the opportunity to use the device of their choice. Thus, mobile devices untether test-takers from the wall in terms of internet access,[1] giving them more degrees of freedom in terms of where they can complete employment-related tests and assessments. However, in spite, or maybe because of these rapid changes, there is limited theoretically and empirically based guidance on the effects of technology on employment-related test and assessment outcomes (Arthur, Keiser, & Doverspike, 2017). Consequently, there has been a call for theoretical and conceptual models and frameworks that formally speak to when one should or should not expect technology-based effects on measurement-related outcomes of interest (Morelli et al., 2017). Furthermore, although most work in this area has focused on personnel selection and employment-related assessment, the issues of interest are potentially relevant to all researchers and practitioners who use UIT assessments in their research and practice.

So, in an effort to address the dearth of theoretical work in this domain, with an emphasis on the technology-mediated devices used to complete UITs, Arthur, Keiser et al. (2017) built on two dimensions—the structural characteristics of UIT devices, and the associated information processing variables that they engender—to develop a framework, the Structural Characteristics/Information Processing

* Corresponding author at: Department of Psychological and Brain Sciences, Texas A&M University, 4235 TAMU, College Station, TX 77843-4235, United States.
E-mail address: w-arthur@tamu.edu (W. Arthur).

[1] Although desktop computers can have wireless internet cards that technically untether them from the wall, Arthur et al. (2017) nevertheless still consider them to be fixed-location devices.

https://doi.org/10.1016/j.intell.2018.02.001
Received 23 July 2017; Received in revised form 9 January 2018; Accepted 8 February 2018
Available online 20 February 2018
0160-2896/

(SCIP) framework, as means for *psychologically* conceptualizing the effect of UIT device-type on test and assessment scores. Arthur et al. (2017) defined a UIT device as "any device that a test-taker can use to complete an unproctored Internet test or assessment where by definition, the test-taker also decides when and where to complete the assessment or test. Thus, a 'UIT device' is not synonymous with a smartphone or other mobile devices. A smartphone or other mobile devices are just one example of a UIT device; so is a desktop computer." (p. 1). Subsequently, they argued that the technological characteristic of being unplugged from the wall ("mobile") versus being plugged into the wall ("nonmobile") failed to provide any conceptual basis or psychologically grounded explanations for why the use of 'mobile' versus 'nonmobile' UIT devices should or should not have an effect on test scores. Thus, the SCIP framework focuses on the psychological aspects of UIT devices that influence test scores.

The basic treatise of the SCIP framework is that differences in specified structural characteristics engender concomitant associated information processing demands, resulting in additional *construct-irrelevant cognitive load* which interacts with the device type, resulting in differential outcomes as a function of the construct (cognitive versus noncognitive) assessed. Hence, Arthur et al. (2017) advanced the SCIP framework as a means of psychologically conceptualizing the effect of UIT devices on assessment and test scores, and subsequently, also as a means for classifying UIT device-types on a continuum of construct-irrelevant cognitive load. Based on a review and integration of the literature, Arthur et al. also demonstrated that the SCIP framework conceptually explains and accounts for the measurement outcome-related findings (e.g., measurement equivalence, mean score differences, criterion-related validity, test-taker reactions and preferences) observed in the literature. They also advanced several testable propositions pertaining to when one might and might not obtain UIT device-type effects on employment-related assessment and test scores. However, there have not yet been any empirical tests of the tenets of the model to date. Consequently, the present study tests one major component of the model, namely, propositions related to working memory.

2. Brief overview of the structural characteristics/information processing (SCIP) framework

Based on the results of a detailed review of the empirical literature, Arthur et al. (2017) identified four structural characteristics that typify *current prototypical* UIT device-types such as desktop computers, notebooks, tablets, and smartphones as *exemplars*. These characteristics are (1) screen size, (2) screen clutter, (3) response interface, and (4) permissibility. Arthur et al. further posited that these four structural characteristics engender the role of four corresponding information processing variables such that (1) screen size → engenders working memory demands, (2) screen clutter → perceptual speed and visual acuity demands, (3) response interface → psychomotor ability demands, and (4) permissibility [distraction] → selective attention demands. So, to the extent that these information processing variables play a role in using the UIT device, they then result in additional *construct-irrelevant* cognitive load that taxes limited attentional resources which, in turn, influences performance on the test or assessment when said cognitive variables or associated demands are not the focal construct of interest (e.g., see Arthur et al., 2014). That is, if one views cognitive resources as finite, then one would expect the addition of cognitive demands irrelevant to the task at hand to detract from one's ability to perform optimally on the focal task. This is illustrated in the general information processing model presented in Fig. 1 which is adapted from Arthur et al. (2017). As Fig. 1 indicates, the SCIP framework conceptualizes assessment device-type effects in terms of how individual differences in the specified information processing variables engendered by the assessment device's structural characteristics intersect with the constructs assessed (cognitive versus noncognitive) to manifest as device-type effects (or lack thereof). Specifically, the structural characteristics stipulated in the model create an information processing-demands continuum, illustrated in Fig. 2, which allows for more nuanced and predictable relationships concerning score or psychometric differences across devices.

In summary, the SCIP framework presents a conceptual explanation for the observed relationships between UIT device types and score differences or lack thereof on cognitive and noncognitive assessments, and permits the formulation of testable propositions as well (Arthur et al., 2017; Morelli et al., 2017). However, there have not yet been any empirical tests of the tenets of the model; thus, as previously noted, the present study tests one major component of the model, namely, propositions related to working memory and screen size.

2.1. Working memory and screen size

The general effects of screen size—the size of the viewable surface on which information is presented—have been examined in the computer-human interaction literature (e.g., Chae & Kim, 2004; De Bruijn, De Mul, & Van Oostendorp, 1992). In terms of the SCIP framework, the role of working memory (Baddeley, 2012) is based on the treatise that differences in screen size (small to large) engender differences in working memory demands (high to low, respectively; see Fig. 2), because more scrolling or screens are typically needed to present the required information and, therefore, more information has to be held in working memory in order to complete the assessment (Sanchez & Goolsbee, 2010).

Working memory has been defined as "a brain system that provides temporary storage and manipulation of the information necessary for … complex cognitive tasks" (Baddeley, 1992, p. 556). In the context of testing, working memory capacity facilitates solving cognitive problems by enabling the test-taker to draw comparisons between alternatives and responses, and then select a response by simultaneously matching the demands of the item, prior knowledge, and relevant information retrieved from memory (Baddeley, 2012). However, in the case of small-screen devices, to the extent that there is limited or only partial information in the test-taker's visual field (i.e., the screen), then working memory will play a greater role in processing the information required to complete the item because larger amounts of it (e.g., portions of current items—either stem or response options; responses to previous similar items) have to be retrieved from working memory (e.g., see Sanchez & Branaghan, 2011). Furthermore, on assessments optimized for small-screen devices such as smartphones, it is not uncommon to present only one item per screen and so the test-taker does not have ready access to their responses to previous items which may inform their responses to subsequent items (e.g., Couper, Conrad, & Tourangeau, 2007; Rivers, Meade, & Fuller, 2009; Schwarz, 1999).

In summary, as per the focus of the present study, the SCIP framework posits that as prototypical or exemplar UIT devices, smartphones are at the high end of the device-engendered construct-irrelevant cognitive load continuum because they are characterized by higher working memory demands due to their smaller screens; in contrast, desktop computers are at the lower end of the continuum because they are characterized by lower working memory demands due to their larger screens.

3. The present study

Arthur et al. (2017) advanced a number of testable propositions that follow from the tenets of the SCIP framework; the present study tests six of them. These propositions pertain to device-type test performance differences, and the relationships between the information processing variables, in this particular instance, working memory, and completion