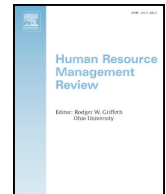


Contents lists available at [ScienceDirect](#)

Human Resource Management Review

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

Take two: Sources and deterrents of score change in employment retesting

Jason G. Randall ^{a,*}, Anton J. Villado ^b

^a Department of Psychology, University at Albany, SUNY, Social Science 387, 1400 Washington Ave., Albany, NY 12222, USA

^b Department of Psychology, Rice University, 6100 Main St., MS-25, Houston, TX 77005, USA

ARTICLE INFO

Article history:

Received 11 September 2014

Received in revised form 10 October 2016

Accepted 11 October 2016

Available online xxx

Keywords:

Personnel selection

Employment tests

Retesting

Practice effects

Individual differences

ABSTRACT

Retesting occurs at all stages of the candidate life cycle—from screening and selection through development and promotion—and therefore has numerous human resource implications. Existing frameworks of retest performance address statistical and conceptual issues related to retest score change; however, little is understood about what causes score gains due to repeat administrations and the effect this has on construct and criterion-related validity. This article reviews the literature on retest effects and introduces a number of factors, including a) construct-relevant change, b) construct-irrelevant change due to individual difference traits, states, and motivation, and c) methods and procedures that may deter construct-irrelevant contamination between tests. In cases where empirical evidence is lacking, propositions are presented to guide future research in evaluating the extent to which construct-irrelevant individual differences and motives contaminate retest performance. We also review ways organizations may limit construct-irrelevant score gains in retesting by attempting to standardize test administration and enhance test security. We end with a discussion of the framework's implications on retest score-based decisions, and methods to mitigate retest score contamination. It is hoped that this paper will facilitate a more comprehensive understanding of retest effects in organizational settings and a way forward for both research and practice.

© 2016 Elsevier Inc. All rights reserved.

Standardized tests administered in education and employment settings hold considerable weight in decision making. Whether a test-taker is testing to gain entry into an organization, to earn a promotion, or to enter a training program, test scores may profoundly influence an individual's career and life. It is not uncommon for an individual to be tested multiple times during a single job search, and many more times across their career. In many of these settings, provisions for re-taking tests are provided (Hausknecht, Halpert, Di Paolo, & Gerrard, 2007; Raymond, Neustel, & Anderson, 2007). The U.S. Department of Labor (DOL; 1999), the American Psychological Association (APA; 1999), and the Society for Industrial/Organizational Psychology (SIOP; 2003) all suggest that when employment decisions are made on the basis of test scores, opportunities to retest should be provided. However, organizations are left to determine how to interpret and use repeat test scores. This determination is further complicated when the retest scores differ from the initial scores, a situation which is likely to be true. Given the prevalence and reliance on test scores for employment decision making, and the likelihood of an individual retesting at some point in their career, understanding the causes and consequences of retest score change is imperative to developing sound retesting policy.

In this paper we introduce a framework of retest score change that identifies several potential sources of change. In doing so, we review and summarize relevant literature concerning individual differences, motivational considerations, constructs, methods,

* Corresponding author.

E-mail addresses: jgrandall@albany.edu (J.G. Randall), antonvillado@rice.edu (A.J. Villado).

and procedures that may be expected to influence retest scores, and where necessary, provide recommendations to deter unwanted score change during retest performance. This effort also clarifies when such factors are most likely to influence test scores and how they might influence construct and criterion-related validity. It is hoped that this review will provide an agenda for future research and a more comprehensive understanding of the practice of retesting in human resource management.

Retest effects

A voluminous literature supports the phenomenon known as the *retest effect* (sometimes also referred to as *testing effect* or *retest bias*). The retest effect refers to score change on a test that arises from merely having taken the test previously, and is well documented in the experimental and educational psychology literatures (Lievens, Buyse, & Sackett, 2005; Roediger & Karpicke, 2006). The direction of the score change is typically positive, although in some paradigms the direction may be negative (e.g., skill decay research). Meta-analytic evidence suggests an increase of about one quarter of a standard deviation between the first and second attempts on measures of cognitive ability (i.e., intelligence or general mental ability [GMA]), with a tapering increase of one fifth of a standard deviation between the second and third attempt (Hausknecht et al., 2007). Similarly, empirical evidence supports the idea that other constructs (e.g., job knowledge, personality), and assessment methods (e.g., situational judgment test, interview) also exhibit score increases as a result of retesting (Hausknecht, 2010; Lievens et al., 2005; Sackett, Burris, & Ryan, 1989; Schleicher, Van Iddekinge, Morgeson, & Campion, 2010; Van Iddekinge, Morgeson, Schleicher, & Campion, 2011; Villado, Randall, & Zimmer, 2016).

The presence of the retest effect has been a challenge both practically, as organizations attempt to understand and interpret multiple test scores, and theoretically, as researchers explore the mechanisms underlying score change. These challenges are relevant for multiple constituents and across many different HRM functions. Test developers and policy makers need to know what they can do to deter unwanted variation in scores between multiple assessments insofar as retesting could harm score reliability and validity. Selection professionals want to know how, when, and why to allow applicants to retest, how to reconcile multiple scores, and if the validity of their assessments is compromised by repeated use. Training professionals have long been interested in what score changes on learning assessments represent (e.g., alpha, beta, gamma change; Golembiewski, Billingsley, & Yeager, 1976).

Retesting research has primarily focused on methodological concerns—providing guidance for statistically evaluating retest score change—and less focused on identifying the sources of these effects. In other words, evidence suggests that scores *will* likely change, and existing frameworks address *how* to detect such changes. However, research addressing *why* scores change is limited. The need to identify the underlying mechanisms of retest score change is critical for two reasons: a) it will help scientists and practitioners develop assessments that might limit the sources of the retest effect, and b) it will inform retesting policies to prevent undesirable score change.

Our goal in this review is to provide a theoretically-grounded and practically-oriented framework for the practice of retesting that comprehensively considers the factors contributing to retest effects, including features of the construct, individual, motivational context, and testing procedure that influence retest performance. We do so by building upon existing frameworks for understanding retest effects to more clearly identify the sources contributing to retest score change. Our approach yields research propositions for future research to address untested, but likely sources of score change, as well as summary points for empirically supported findings. The organization of the literature around a simple, yet comprehensive model adds value to the literature by providing a means to organize the extant research, as well as a means of introducing issues one must consider when formulating a retest policy.

Such an endeavor requires a consideration of the multiple uses of retesting in organizational settings (e.g., selection, training, screening, promotion) as well as the multiple constructs and methods used, so we attempt to be sensitive to this by providing evidence and examples from various sources, separating constructs from methods (Arthur & Villado, 2008; Binning & Barrett, 1989) and by incorporating specific propositions and summary statements to address these issues. Certain constructs have received more attention than others in the domain of retesting in employment settings. Specifically, cognitive ability testing is the predominant focus of the research on retesting, and consequently, a greater preponderance of the literature considered in this review reflects this emphasis. There is also a growing literature focused on personality faking, therefore, we also draw upon this research to inform our review and model of employment retesting. Additionally, research on retesting has been conducted in employment, education, and laboratory settings, and although we draw upon research in all areas to inform our framework and propositions, we focus our implications, conclusions, and future research directions more specifically to employment settings. This focus helps address our primary goal of providing a comprehensive framework of the theoretical causes of retest effects to guide future research and inform practice.

Previous frameworks for retesting

Historically, the issue of accounting for multiple test scores has typically taken a statistical and methodological approach. Lievens, Buyse, and Sackett (2005) proposed and tested a framework to help interpret score change at the within- and between-person level. This framework provides guidance to answer key questions about retesting in operational settings, including: how to detect individual score change over time (paired *t*-test, within-person *d*), how to evaluate whether the initial or retest score has higher criterion-related validity (*z*-test of dependent correlations), and whether the criterion-related validity of test scores differs between one-time and repeat test-takers (differential validity/prediction). The utility of this framework (Lievens

Download English Version:

<https://daneshyari.com/en/article/5033634>

Download Persian Version:

<https://daneshyari.com/article/5033634>

[Daneshyari.com](https://daneshyari.com)