American Psychological Association

Journal of Applied Psychology

0021-9010/19/$12.00

2019, Vol. 104, No. 9, 1089–1102
http://dx.doi.org/10.1037/apl0000401

Using Practice Employment Tests to Improve Recruitment and Personnel Selection Outcomes for Organizations and Job Seekers

Michael C. Campion
University of Texas Rio Grande Valley

Emily D. Campion
Old Dominion University

Michael A. Campion
Purdue University

This study introduces the use of practice employment tests during recruitment as a tool with the potential to improve outcomes for both an organization and its (potential) applicants during personnel selection. Synthesizing research on recruitment, selection, job search, adverse impact, signaling theory, and human capital theory, we propose that practice tests reduce information asymmetry regarding the nature of an organization's assessment procedures, thereby acting as short-term human capital investment opportunities. Using a large sample of potential applicants and applicants who later decided to apply for jobs within a professional occupation in a large organization, we demonstrate that (a) those who took the practice tests scored higher on the actual tests; (b) score gains between practice tests and actual tests were greater for Blacks and Hispanics when compared to Whites; (c) the practice test exhibited a self-selection effect, encouraging those with higher scores to apply; and (d) score gains between practice tests and actual tests were similar to scores observed for those retesting on the actual tests. These findings suggest practice tests may be capable of simultaneously enhancing organizational outcomes (e.g., increased quality of applicants, reduced cost of testing unqualified applicants, and reduced adverse impact) and applicant outcomes (e.g., increased human capital, increased chances of eventual employment, and reduced disappointment and wasted effort from unsuccessful application).

Keywords: recruitment, personnel selection, job search, staffing, human capital

The most important goal of recruitment for jobs in professional occupations is not to attract a large number of candidates. It is, instead, to attract qualified candidates. However, recruiting efforts in many organizations for these types of jobs are, sometimes inadvertently, focused on accumulating a large number of candidates regardless of their qualifications. In fact, it is common for organizations to define recruiting success in part by how many candidates they attract (Dineen & Williamson, 2012). The problem is that some candidates who have little chance of being selected, thus creating a quality−quantity dilemma for organizations, are recruited. This problem has been exacerbated in recent times by the use of Internet recruiting, because it reaches larger numbers of potential candidates and because it is usually accompanied by online applications making it easy to apply (Cappelli, 2001; Dineen & Noe, 2009).

Recruiting a large number of candidates, regardless of their qualifications, creates many problems for organizations. First, it requires good selection procedures to separate candidates who are qualified from those who are unqualified. Second, it increases the costs of selection because more candidates must be screened. Third, it may increase adverse impact against protected groups of candidates if there are any subgroup differences on the selection procedures, because larger samples make differences more likely to be statistically significant. The larger number of rejected candidates also increases the financial liability should an organization be unable to defend its selection procedures in a lawsuit. Unfortunately, the outreach efforts of many organizations to recruit racially and ethnically diverse candidates can aggravate this problem if they yield larger numbers of such candidates who fail the selection procedures (Campion, 2015; Newman & Lyon, 2009).

At the same time, potential applicants are interested in becoming more qualified and identifying where they should concentrate their application efforts to realize the greatest payoff (Tomlinson, 2007, 2008). For example, recent research has suggested potential applicants are interested in making *long-term* investments to enhance their quality and that an organization may be able to help them with this by providing more diagnostic information early on in their educational careers (e.g., high school; Campion, Ployhart, & Campion, 2017). However, potential applicants are likely to be

This article was published Online First March 14, 2019.

Michael C. Campion, Vackar College of Business and Entrepreneurship, University of Texas Rio Grande Valley; Emily D. Campion, Strome College of Business, Old Dominion University; Michael A. Campion, Krannert School of Business, Purdue University.

An earlier version of this study was presented at the annual conference for the Society for Industrial and Organizational Psychology in Orlando, Florida, in 2017.

Correspondence concerning this article should be addressed to Michael C. Campion, Vackar College of Business and Entrepreneurship, University of Texas Rio Grande Valley, 307 East Railroad Street, Weslaco, TX 78596. E-mail: michael.campion@utrgv.edu

especially interested in, and willing to make, *short-term* investments directly prior to application to enhance their quality and gather information regarding whether they should apply. Organizations may be willing to help potential applicants achieve these outcomes through their recruitment efforts in order to realize important additional benefits of their own, such as increased candidate preparation for the hiring process, encouragement of strong candidates to apply, and reductions in subgroup differences during personnel selection. Thus, the question becomes: What type of practice might jointly satisfy these needs of both the organization and its potential applicants in the short term?

The purpose of this article is to advance staffing scholarship by synthesizing research on recruitment, selection, job search, adverse impact, signaling theory, and human capital theory to examine a new approach that has not been previously studied: the provision of a practice test that prospective candidates can take to determine their likelihood of being selected if they apply. Organizations that use testing for academic admissions, licensure, professional credentialing, and related purposes have long provided study guides and practice test questions (Powers, 1986). Some organizations that use testing for hiring will offer these materials as well, but it is less common. Providing such information is usually considered useful because it helps to prepare individuals for the tests and to improve their reactions to testing. Thus, the unexplored, but theoretically and practically important, issue is whether and why the use of practice tests can improve outcomes of value to organizations that use them for hiring (i.e., increased applicant quality, reductions in applicant pool quantity and potential for adverse impact) as well as their potential applicants (i.e., enhanced human capital in the short term, application decisions).

In this study, we introduce the concept of using practice tests as a means of mitigating the candidate quantity−quality dilemma and contribute to the literatures on staffing and job search in at least four ways. First, we develop and test theoretical predictions proposing that practice tests serve as signals for potential applicants, which can then be leveraged to enhance their quality prior to application. In doing so, we demonstrate that scores on the actual test are higher for applicants who took the practice test when compared to applicants who did not. Previous research has suggested that recruitment sources signal information to applicants regarding how to improve their quality in the long term (e.g., educational choices; Campion et al., 2017). Unfortunately, such long-term solutions are irrelevant to potential applicants who are nearing or already on the job market. Thus, what is needed is an extension of this theory that examines complementary practices, such as the use of practice tests, that can be adopted that result in more immediate benefits to organizations and their potential applicants.

Second, we provide initial evidence suggesting that practice tests may be a useful tool toward mitigating the adverse impact−validity dilemma (Ployhart & Holtz, 2008). Specifically, we show that, on average, score gains between the practice and actual tests are greater for traditionally disadvantaged minority groups (i.e., Blacks and Hispanics) when compared to Whites. Previous research examining ways to contend with this dilemma has generally focused on the types of selection procedures used (Ployhart & Holtz, 2008) and/or the characteristics of the selection process (e.g., ordering−weighting−adding predictors; Sackett & Ellingson, 1997; Sackett & Roth, 1996). Other previous research has even begun to consider whether alternative ways of scoring procedures can reduce adverse impact (e.g., using text analytics and predictive modeling computer software programs; Campion, Campion, Campion, & Reider, 2016). Our findings suggest that it may also be useful to look "upstream" of the selection process to an organization's recruitment practices to consider how they might be used to achieve this goal.

Third, we extend prior research on recruitment methods by explaining how the communication of realistic information regarding an organization's selection criteria, rather than a job (Breaugh & Starke, 2000; Phillips, 1998), influences the immediate decisions of potential applicants of differing levels of quality. We demonstrate that individuals who score higher on the practice test are more likely to apply. By reducing information asymmetry regarding the difficulty of an organization's assessments and offering potential applicants the opportunity to perform, practice tests provide a way for an organization to communicate valuable signals to potential applicants that enable them to make more informed decisions regarding whether to apply. Further, by potentially reducing the inadvertent legal risk associated with affirmative−outreach recruiting and its tendency to increase the number of unqualified minorities who apply (Newman & Lyon, 2009), practice tests may actually encourage organizations to conduct more affirmative recruiting of minorities.

Finally, we provide initial evidence demonstrating that, overall, the effects of practice testing and retesting are equivalent. Practically speaking, this is an important finding because offering practice tests is often far cheaper than offering the ability to retest. Thus, it may serve a cost-cutting function for organizations (and applicants). Theoretically, we propose that the options of retesting and taking practice tests both act as signals because they allow an organization to reduce information asymmetry regarding the nature and difficulty of its assessment procedures. However, retesting is much more costly to an organization than is offering practice tests. Thus, by identifying when and where a less costly signal might achieve comparable results and thus be more valuable than a more costly signal in terms of payoff to an organization (and its applicants), we refine the boundary conditions of signaling theory as it applies to the context of staffing in organizations (Bangerter, Roulin, & König, 2012).

Theoretical Background and Hypotheses

Description of Practice Tests

Prior to our review of relevant theory and development of our hypotheses, it is important to first briefly outline the general format of the practice employment test examined in this study. The practice test is a retired version of a previous test used by the organization, and it contains two subtests that assess occupation-specific knowledge and English writing skills. The practice test was made available online to potential applicants on the organization's website, includes 50 knowledge items and 45 English skills items, and takes approximately 70 min to complete. The practice test is not a condition of application, for then it would be considered a selection procedure. Instead, it is a test preparation tool that potential applicants can take voluntarily and anonymously and only they receive their scores and information regarding their likelihood of passing the organization's actual assessments based

on normative data. They also have the opportunity to review the correct answers to each question. Finally, it is accompanied by information offering additional preparation guidance (e.g., suggested classes or readings) that is not tied to individuals' scores.

Key Concepts From Signaling and Human Capital Theories

Signaling and human capital theories provide a useful foundation for understanding how practice tests operate and why they might improve outcomes for organizations and applicants during personnel selection. Signaling theory's focus is on why cooperative behavior occurs among parties (e.g., organizations and applicants) who have both limited information regarding the other and at least partially misaligned goals (Bangerter et al., 2012; Connelly, Certo, Ireland, & Reutzel, 2011; Rynes, Bretz, & Gerhart, 1991). This theory argues that *cooperative behavior,* or joint action with the goal of achieving mutually beneficial outcomes, is made possible by costly signals. More costly signals prompt such behavior because, when compared to less costly signals, they communicate more reliable information from one party to another. As such, they are considered advantageous for at least two reasons. First, costly signals are beneficial to signal senders (e.g., organizations) because they afford advantages over their competitors (e.g., attracting higher quality talent, increasing fairness perceptions regarding selection procedures). Second, costly signals are beneficial to signal receivers (e.g., job seekers) not only because they tend to more directly communicate unobservable qualities of signal senders but also because they reduce search costs they must incur to otherwise acquire useful information (Spence, 2002; Stiglitz, 2000). For example, Bangerter et al. (2012) proposed that, in personnel selection contexts, job seekers expend effort (invest) in an attempt to identify the organization's selection criteria and use this information to their advantage when seeking jobs. In the context of this study, one example of a costly signal is retesting because the organization incurs relatively high costs when it allows applicants to retake the entrance exams. In contrast, by requiring only a relatively low, one-time setup cost, practice tests serve as a less costly signal. Such lower cost signals are important to examine in general and in comparison to higher cost signals, because they may operate similarly in some contexts (i.e., through enabling cooperation between two opposing parties) and also result in enhanced outcomes of value to both an organization and its (potential) applicants during recruitment and selection while simultaneously reducing costs.

Human capital theory extends this notion by identifying and explaining the ways in which individuals translate organizational signals into economically valuable outcomes for themselves. This theory argues that individuals invest in themselves by making decisions about their education, experience, and information-gathering efforts (Becker, 1962, 1964; Strober, 1990). Investment decisions with high rates of return (e.g., advanced education, work experiences) augment individuals' human capital because they can increase productivity and wages (Becker, 1964; Strober, 1990). It is important to recognize that time and information are essential to human capital theory. First, there is often a general distinction made between long- and short-term investments. Whereas long-term investments are primarily viewed as enabling greater amounts of knowledge and skill accumulation, short-term investments also provide valuable market information. For example, Becker (1962) suggested that information regarding the economic system as it relates to organizational requirements for different knowledge, skills, abilities, and other characteristics (KSAOs), increases job seekers' command over their resources (e.g., time, effort, KSAO levels). Second, the timing of decisions matters inasmuch as it may constrain whether long-term or only short-term investments are possible. For example, the timing of application decisions influences the amount of time applicants have had to make investments with the potential to enhance their human capital. Third, information regarding the requirements for obtaining a job within a given occupation is important because it helps to guide individuals' investment decisions toward those likely to produce the greatest amount of return given their investment time frame.

Practice Employment Tests and Actual Test Scores

Whereas Bangerter et al. (2012) proposed that organizations continually search for ways to "keep their selection criteria from being identified" (p. 727), we propose that there are certain advantages associated with doing the opposite. We begin by discussing why taking practice tests may be associated with improved actual test scores during personnel selection. First, practice tests and their associated feedback may enable an organization to help applicants improve more traditional forms of human capital prior to application. Receiving information regarding the content of an organization's assessments enables individuals to focus their short-term investment efforts on acquiring knowledge and skills on topics that are relevant to obtaining a job within the occupation at the organization (Lievens, Buyse, & Sackett, 2005; Messick & Jungeblut, 1981; Powers, 1986; Sackett, Burris, & Ryan, 1989). For example, assessment information may serve as a general study guide that identifies what types of knowledge and skills areas an organization demands as well as the levels of knowledge and skills required.

Second, practice tests and associated feedback may enable an organization to communicate information useful to applicants' human capital investment strategies for improving other characteristics aside from those typically examined in human capital research that are nonetheless critical to obtaining a job. For example, by offering potential applicants the opportunity to perform, practice tests may improve test-taking strategies (Hausknecht, Day, & Thomas, 2004; Sackett et al., 1989). Similarly, prior research has suggested that practice testing may reduce test-taking anxiety (Anastasi, 1981; Hausknecht, Halpert, Di Paolo, & Moriarty Gerrard, 2007; Messick & Jungeblut, 1981) and improve test-taking self-efficacy, test-taking motivation, and beliefs regarding test validity (Schleicher, Van Iddekinge, Morgeson, & Campion, 2010). Personnel selection research typically views these factors as influencing systematic (or constant; see Nunnally & Bernstein, 1994) error, which depresses scores on selection procedures (Lievens et al., 2005). Although this is likely the case, there is an advantage to conceptualizing these factors as components of human capital. Doing so implies applicants, at least in this context, can play a more proactive role in mitigating systematic error upon initial assessment so that this responsibility is not left entirely to the organization (e.g., through use of fairness-enhancing techniques), the outcome of which benefits both the organization and its applicants.

Finally, practice tests may enable organizations to augment the accessibility of knowledge and skills relevant to performance on actual tests (i.e., human capital) for applicants. *Accessibility* refers to the degree to which knowledge, for example, is readily available in memory (Feldman & Lynch, 1988; also see Van Hoye & Lievens, 2009). Knowledge that is more accessible is more likely to influence judgments, decisions, and behavior. For example, research on syntax and sentence formulation has suggested that the accessibility of lexical concepts influences sentence formulation, which may be relevant in testing for English skills (Bock, 1982). Research on memory and cognition has also suggested that accessibility of knowledge categories influences accuracy of recognition as well as time required to identify objects, which may be relevant to multiple-choice tests of job-relevant knowledge (Bargh, Lombardi, & Higgins, 1988; Bruner, 1957). Finally, research on learning through the use of tests has shown that taking a test while learning and relearning concepts may slow down the process of forgetting information (Rawson & Dunlosky, 2011; Roediger & Karpicke, 2006). In this way, taking a practice test and receiving feedback can be viewed as a way for applicants to refamiliarize themselves with knowledge and skills relevant to actual test performance that they had previously learned but had forgotten. It may also attenuate forgetting effects in the short term. For these reasons, we hypothesized the following:

> *Hypothesis 1:* Applicants who take the practice tests will have higher scores on the actual tests than will applicants who did not take the practice tests.

Practice Employment Tests and Subgroup Differences

An important concern in personnel selection is whether assessments have adverse impact. *Adverse impact* refers to a situation where passing rates for minority candidates on an assessment (e.g., Blacks, Hispanics, female applicants) are lower than are passing rates for the majority subgroup (e.g., Whites, male applicants; Uniform Guidelines, 1978). Prior research on personnel selection has identified a number of ways to potentially mitigate adverse impact during assessment (e.g., Campion et al., 2016; Ployhart & Holtz, 2008). However, this research has generally ignored whether recruitment practices could also be used for this purpose (for an exception see Newman & Lyon, 2009). Thus, it is of interest to explore whether practice tests might improve or exacerbate mean differences in candidate scores.

On the one hand, research on retesting has suggested practice tests may result in greater score gains for nonminorities. For example, Schleicher et al. (2010) and Van Iddekinge, Morgeson, Schleicher, and Campion (2011) both found that Whites tended to realize greater scores gains from retesting than did Blacks and Hispanics on both English skills and knowledge tests. Potential explanations offered for these findings include differences in scores on mental ability tests, which may reflect one's ability to identify, learn, and use test-related information from the first test to the next (also see Kulik, Kulik, & Bangert, 1984), and the possibility that those holding more positive test-taking attitudes and greater motivation would gain more because they believe their test scores reflect competence and are within their control.

On the other hand, other research has suggested minorities may realize greater score gains from practice tests. Prior research has suggested minorities tend to hold more negative test-taking attitudes (e.g., test self-efficacy, test-taking motivation; Chan, 1997; Ryan, 2001). They may also experience stereotype threat when taking employment tests (Nguyen & Ryan, 2008; but cf. Zigerell, 2017). Therefore, in general, minorities may have more room for improvement on these characteristics that influence scores. Because they can be taken remotely and anonymously, practice tests offer a low-stakes context in which these characteristics can be proactively improved prior to taking the actual employment tests. Supporting this, research has suggested practices that offer the opportunity to perform, such as practice tests, are critical to improving candidate reactions, especially for those who perform less well on assessments (Schleicher, Venkataramani, Morgeson, & Campion, 2006). This should reduce any systematic error that is depressing minority scores on actual tests to a greater degree, resulting in greater gains for these subgroups. Given theoretical support for both assertions, we examined the following:

> *Research Question:* Will within-subgroup gains in scores between the practice tests and the actual tests be greater for minorities or for nonminorities?

Practice Tests and Application Decisions

In addition to improving the quality of an organization's applicant pool through enhancing applicants' human capital prior to application, practice tests may also improve the quality of an organization's applicant pool by encouraging potential applicants to apply who already have high levels of human capital. Prior research on recruitment sources and realistic job previews touches on this notion. Unfortunately, although research on these topics has supported effects on turnover, it has not consistently supported the predictions that realistic information affects attraction or attrition from the recruitment and selection processes (Barber, 1998; Breaugh & Starke, 2000; Phillips, 1998; Zottoli & Wanous, 2000). However, this research focused on recruitment methods conveying information regarding jobs not selection procedures.

Signaling theory suggests that reducing information asymmetry for potential applicants regarding the nature and difficulty of the organization's assessments and their probability of passing is likely to affect potential applicants' decisions to apply. Such signals communicate information necessary for potential applicants to focus their job-seeking time and effort by enabling the development of hiring expectancies (or perceptions regarding their likelihood of obtaining a job within the occupation). This is important because the assessment procedures typically used during personnel selection for jobs within professional occupations tend to be very time-intensive and draining for applicants (e.g., multiple or long assessments, multiple interviews, travel). As such, potential applicants' decisions would benefit greatly from information regarding their current levels of human capital as well as their likelihood of passing.

Previous research has supported the value of such signals to potential applicants. For example, research has suggested that this information is often sought out where possible by potential applicants to gain a competitive edge while minimizing effort expended (Bangerter et al., 2012). Research has also consistently found hiring expectancies to predict levels of attraction to organizations (Uggerslev, Fassina, & Kraichy, 2012). In addition to reinforcing

the application intentions of those taking practice tests, practice tests may also encourage application of those who are hesitant to do so. For these reasons, we hypothesize the following:

Hypothesis 2: Those who score higher on practice tests will be more likely to apply.

Practice Employment Test Versus Retesting Gains

According to signaling and human capital theories, the costs associated with emitting signals and accumulating human capital are often just as important as are the actual information communicated and the amount of human capital accrued. This is because the costs involved in signaling and investing in human capital limit which options are probable or even possible. Typically, the greater the costs involved, the greater the payoff should be (Bangerter et al., 2012; Becker, 1964; Connelly et al., 2011).

Offering practice tests and offering the option to retest represent two types of organizational signals to (potential) applicants in that both reduce information asymmetry regarding the nature and difficulty of the organization's assessment procedures. However, these two types of signals differ in cost for the organization and its applicants. They may also differ in terms of the extent to which they lead to mutually beneficial outcomes, such as increases in applicant quality. Retesting involves an organization's allowing an individual who has failed an assessment to return at a later date and retake it. Retesting is widely allowed in personnel selection as well as in other contexts such as education (e.g., Scholastic Aptitude Tests, graduate management admission tests; Hausknecht et al., 2007; Schleicher et al., 2010). However, unlike educational contexts where the cost of retesting is entirely incurred by the student, in personnel selection contexts the cost of retesting falls on both the applicant and the organization. For example, in 2017 the cost to proctor applicants for organizations at one typical commercial testing center was $45 per hour per seat, which would translate into $90 per candidate in the present context where the test is 2 hr. In the present context, where almost 50%, or 7,000 individuals, retest each year, the total cost would be $630,000. Conversely, designing and implementing the practice tests in the present context required only a one-time expenditure of slightly less than $5,000. Similarly, from the applicant's perspective, retesting involves at least 8 hr of time, which includes transit, registering for assessments, and taking the tests, plus financial cost of transportation. In contrast, practice tests in this context can be taken from home for no additional costs and require approximately 70 min of time. Given these differences in cost, it is important to explore whether these practices yield proportional increases in terms of candidate quality.

Such "lower cost" signals may actually be more valuable to both the organization and its applicants because the human capital gains associated with taking the practice tests may be greater than are those associated with retesting. For example, given the lower stakes involved, practice tests require less effort and are less stressful for the applicant. Thus, applicants may be able to absorb more information regarding the assessments (Rawson & Dunlosky, 2011). For example, experimental research comparing the same activities across low- to high-stakes contexts has shown that memory and recall are better under low-stakes conditions (e.g., using an associations learning task; Ariely, Gneezy, Loewenstein, & Mazar, 2009). Research has also suggested that anxiety associated with performing in educational testing in high-stakes contexts can inhibit the functioning of working memory, which affects one's ability to perform and learn (Ashcraft & Kirk, 2001; Beilock & Carr, 2005). In addition, practice tests, at least in the present context, are accompanied by additional information regarding how potential applicants can improve their scores in the future, and this information would otherwise require more searching on the potential applicant's part to collect it. For these reasons, we hypothesized the following:

Hypothesis 3: The gains between the practice and actual tests will be greater than will the gains associated with retesting.

Method

Participants

The initial sample for this study included 25,548 participants who took the practice tests and 19,089 participants who applied to the organization and took the actual tests. We retained practice test data only if individuals got at least 25% correct (the level of chance) on each test and completed both tests (clicked through all the pages) to generate a final score. We also allowed for up to 50% of items per test to be skipped. These rules approximated the estimated range of legitimate scores for several reasons. First, these rules ensured that the individuals attempted most items, reviewed all the items, and did not respond randomly. Second, some individuals scored this low or lower on the actual tests upon application. Third, a purpose of practice tests is to allow individuals to evaluate whether they should apply, which includes potential applicants who are woefully unprepared and may realize that fact after starting the practice tests and thus not finish. Fourth, opening the exam and exploring, but not answering, more than 50% of the questions may provide speculative individuals a realistic preview of the nature of work in this occupation at this organization, but these scores are not true estimates of their levels of human capital. Elimination of practice test takers based on these rules resulted in a final sample of 23,238 practice test takers. To ensure comparability with the actual tests, we also eliminated 77 individuals who scored below 25% correct (level of chance) on each test (final applicant $N = 19{,}012$). Whereas these eliminations of applicants did not influence the results or their interpretation, they did improve our analyses because they helped to control for differences in test takers' motivation between practice and actual tests.

Within this sample, 6,169 individuals who took both a practice test and the actual test could be matched based on e-mail addresses they provided to the organization. In addition to the practice tests' playing a role in encouraging those likely to be successful to apply and indirectly not encouraging those unlikely to be successful, reductions in sample size occurred due to the need to match candidates and the fact that many people who were not really intending to apply likely took the practice test out of curiosity (e.g., individuals still in high school or searching for jobs in general). Sample characteristics appear in Table 1. Applicants generally all had an undergraduate degree, and 50% had a master's degree.

Individuals applied for one of five entry-level, rotational jobs within an occupation in a U.S. government agency between 2014 and 2015. As part of the entrance exam, all individuals take the

Table 1
Sample Characteristics

Sample	*N*	Race			Gender		
		Black	White	Hispanic	Other	Female	Male
Took practice test	23,238						
Took practice test only	17,069						
Took practice and actual test	6,169						
n		522	4,286	601	760	2,373	3,696
%		8.46	69.48	9.74	12.32	38.47	59.91
Took actual test	19,012						
n		2,090	12,251	2,108	2,640	7,003	11,715
%		10.88	64.44	11.09	13.89	36.83	61.62
Took actual test only	12,843						
n		1,549	7,938	1,494	1,862	4,613	7,962
%		12.06	61.81	11.63	14.50	35.92	62.00
Took retest	451						
n		49	303	38	61	136	302
%		10.86	67.18	8.43	13.53	30.16	66.96

Note. Data on test takers' race and gender were not collected while taking the practice test. *Other* for race includes "did not respond." Gender does not sum to total *N* due to missing data.

same battery of employment tests. Similar to many other professional occupations, this occupation requires a high level of knowledge and skill. However, the likelihood of passing the assessments can also be enhanced through increasing candidates' other characteristics (e.g., test attitudes) and/or enhancing KSAO accessibility. Thus, it offers a context where preparation in the form of short-term human capital investments is important.

The organization would not allow us to identify the occupation or jobs as a condition of publication because this is a current hiring program and the organization does not want to disclose information that may compromise its staffing process. The occupation consists of five job areas that include working in professional areas such as management and economics. Incumbents are generalists who are transferred regularly and expected to work in any of these five job areas depending on staffing needs. A specialty is selected at the time of application, but all recruiting is for the combined set of job areas and the hiring procedures are the same, with roughly equal numbers of individuals hired in each. This organization is a major, but not the only, employer of this occupation, and these jobs include the range of jobs that constitute this occupation.

This organization does a tremendous amount of outreach to increase minority applicants. This results in effectively doubling the number of minority candidates compared to the relevant labor market for these jobs.[1] The recruiters of this organization, who devote most of their attention to minority recruiting, asked for a practice test for two reasons: (a) to provide a way to encourage hesitant candidates to apply and (b) to reduce candidates' disappointment. They felt the practice test would enhance their affirmative-recruiting success.

The data reported in this study were part of a larger data collection effort that has been ongoing since the year 2000. This comprehensive data collection has allowed several contributions to the research literature (Campion et al., 2016, 2017; Levashina, Morgeson, & Campion, 2009, 2012; McCarthy, Van Iddekinge, & Campion, 2010; Schleicher et al., 2010, 2006). However, the data, the sample, and theoretical development of the hypotheses in this study do not overlap with those used in previous studies. This study was submitted to the Purdue University Institutional Review Board and judged to be exempt (Pro1701018681). Sample sizes for measures described later vary downward due to differences in types of information sought from potential and actual applicants as well as their discretion in responding to items (e.g., demographic variables). We therefore report *N*s in all tables because we tried to use the largest sample size possible for each analysis. Statistical power exceeded 90% for all hypothesis tests.

Procedure

The practice tests were described previously. Because the tests are equivalent in style and length, scores achieved reflect an estimate of the scores on the actual tests with good probability. Moreover, because the practice tests were previous actual tests, extensive normative data existed from which to estimate the potential applicants' likelihood of passing the actual tests.

Potential applicants can take the practice tests anytime from the convenience of their personal computers by accessing a link posted on the organization's website. Instructions tell potential applicants to access the link and enter their e-mail address to begin the practice tests. Upon answering a survey, which asks about what other forms of preparation they have engaged in prior to taking the practice test, individuals are allowed to begin the practice tests. Each test is timed, and the timer cannot be stopped once individuals begin. Thus, they are encouraged to allocate sufficient uninterrupted time to take the test. However, if they lose computer connectivity, they can log back on with their e-mail to resume where they left off. There is no help desk service provided with this computer application, although there is a help page with technical information and an e-mail address to report technical problems.

[1] Relevant labor market percentages were based on Bureau of Labor Statistics data on employment in comparable jobs throughout the government, consistent with recommended procedures for determining labor market availability for affirmative action goal setting from the U.S. Department of Labor, Office of Federal Contract Compliance Programs (2014).

After individuals complete the tests, scores appear immediately. They receive feedback on each test item, separate scores on the two practice tests, and the combined score, which is the same way scores are provided for the actual tests. They also receive additional interpretive information, including whether they passed the practice tests and their probability (as a percentage) of passing the actual tests. Because there is a small margin of error, passing the practice tests is no guarantee of passing the actual tests, but the higher the score, the greater the likelihood. In addition, they are provided a list of potential study materials that could help to improve their scores.

The recruiting website and the instructions to the practice test provide the following recommendations to potential candidates regarding how to improve their chances of success. First, the organization recommends they take the practice tests. They are further told that this will give them a reliable indication of how well they might perform on the actual tests and that it will enable them to study or pursue further education before taking the actual exam. They are also told that if they are likely to pass the actual tests, the practice tests will give them the confidence and motivation to move forward with their application. If they are unlikely to pass, the feedback should encourage them to postpone taking the real test until they have improved. Either way, the results should make them better prepared for the exam, which should increase their score and also reduce test anxiety somewhat. Passing scores on the practice tests were set to approximate a 50% passing rate on each exam, which approximates the passing score used by the organization.

Second, individuals are told that they should take the practice tests as though they were real applicants. This means they should prepare as though they were taking the actual tests, take the practice tests one after the other, take the tests at a time when they are alert and prepared, allocate sufficient time to take the entire exam in one sitting, and so on. Third, they are told that if they do not perform as well as they might like on the initial practice tests, then they should study and take the practice tests again. For example, studying might range from reading on the topics covered on the practice tests to taking courses or pursuing degrees related to the work in this occupation. Finally, they are told that taking the practice tests is provided as a service to potential applicants to help them prepare. It does not promise or imply any certainty of passing the actual tests.

If potential applicants took the practice tests multiple times, we used their first scores for our analyses. Only the actual assessments collect information on applicant race and gender, so diversity analyses are limited to those who took the actual tests.

Measures

Actual tests. *Application* was operationalized in this study by coding individuals who had applied to the organization and taken the actual tests as 1 and those who had not as 0. *Scores on the actual tests* was operationalized as scores on a knowledge test and an English writing skills test, which are taken during the first hurdle of the hiring process. The organization developed these tests with the assistance of external consultants. They take approximately 1 hr each to complete (including instruction time and items being pilot-tested). In developing these tests, a large-scale job analysis of each of five job areas constituting this occupation at this organization was performed. These data were converted into a test-specification plan to ensure content validity. For the knowledge test, the specification identifies a specific number of items for each of more than 20 knowledge areas that were identified as common to all five jobs within this occupation at this organization. The English skills test is a job sample requiring applicants to respond to examples of text from the job that correct the grammar, punctuation, structure, word usage, and other aspects of written English. The knowledge test contains approximately 50 multiple-choice items, and the English skills test contains approximately 45 multiple-choice items, which are scored as "correct" or "incorrect." The internal consistency reliability of scores on the knowledge and English skills tests were $\alpha = .92$ and $\alpha = .90$, respectively. The actual tests are intended to be equivalent over time (and thus also with the practice tests because they are retired actual tests) in terms of content (by virtue of using the same specifications) and difficulty (by matching on average item difficulty), and scores on the tests were converted to *T* scores to ensure that the means and standard deviations were the same. The tests are highly intercorrelated ($r = .59$), as might be expected based on past research and because they both tap into mental ability (e.g., Ree & Earles, 1991), and are summed to create a combined score.

Practice tests. *Taking the practice tests* was operationalized by coding individuals who had taken the practice tests as 1 and those who had not as 0. *Scores on the practice tests* were operationalized using scores on retired versions of the knowledge and English writing skills tests as described earlier. The practice tests are shorter (35 min each) because they do not include items being pilot-tested and additional instructions. The internal consistency reliability of scores on the knowledge and English writing skills tests were $\alpha = .92$ and $\alpha = .90$, respectively. As with the actual tests, scores on the practice tests were converted to *T* scores and summed to create a combined score (intercorrelation = .61). The *T* scores were based on the normative data from the actual use of this test and not based on the sample that took the practice test, so that the feedback was based on the pool of actual test takers.

Retesting. To examine the effects of retesting, we received from the organization the data on candidates who had applied and taken the assessments in 2014 for the first time and then retested in 2015 but who had not taken the practice test. Candidates are allowed to retake assessments only once a year. Thus, these candidates had retested only one time. These data were compared with the data on practice test takers, who had the opportunity to take the actual tests only once during the same time span. When the effects of retesting were examined, individuals were coded as 1 if they had taken the actual tests again and 0 if not.

Minorities and nonminorities. For analyses of subgroup differences, minority status by race (i.e., Black, Hispanic) was coded as 1, and nonminority status (i.e., White) was coded as 0. Minority–nonminority comparisons were investigated separately for each group (i.e., Black vs. White, Hispanic vs. White). We focused on these minority subgroups because they are the ones that generally show the greatest differences in testing during personnel selection (Hough, Oswald, & Ployhart, 2001), and thus understanding how to reduce their subgroup differences is a primary concern for organizations and scholarship (Avery, Hernandez, & Hebl, 2004; Newman & Lyon, 2009).

Supplemental analysis variable. Several supplemental analyses were performed to consider alternative explanations for our

findings. In performing these analyses, we examined *biodata* scores for applicants, which are another component of the hiring process at this organization. The biodata questionnaire was developed based on a job analysis to measure planfulness, motivation, initiative, and other dimensions that may influence or reflect applicants' quality. The items asked for self-report descriptions of past experiences. Examples include the following: How often have you used a budget to plan your expenditures? How often have you set goals for yourself, which you knew from the beginning would be difficult for you to achieve? How many times have you had to assume an authority role to improve a group's efficiency? It contains 56 items, was scored on a 5-point scale, and were *T*-scored ($\alpha = .95$). The 5-point responses were summed to create the total scores but were keyed based on correlations with the next stage of the hiring process, which was a structured three-person panel rating of experience and education information. The biodata scores correlated $r = .14$ ($p < .01$) with the knowledge test and $r = .17$ ($p < .01$) with the English writing skills test.

Results

Hypothesis Tests

Table 2 presents the means, standard deviations, and intercorrelations among the study variables. Because the tests are *T*-scored, the means tend to approximate 50 and the standard deviations 10, with the values for the composites about twice that. Of special interest are the correlations between the initial practice tests and the actual tests, which is .65 for job knowledge, .70 for English skills, and .77 for the composite. They can be viewed as a form of alternative forms reliability, or, more relevant to the applicants, an indication that the practice test scores are good estimates of actual test scores.

Hypothesis 1 predicted that applicants who took the practice tests would have higher scores on the actual tests than would applicants who did not take the practice tests. As Table 3 indicates, this hypothesis was supported for total scores on the combined actual tests ($d = .27$) as well as for each test individually ($d = .19$ for the knowledge test and $d = .29$ for the English skills test). In practical terms, assuming an average of 15,000 candidates per year and a 50% passing rate, this is equivalent to an increase in passing rates of approximately 11%, which translates into about 825 additional applicants' being capable of passing the assessments per year.

Table 3
T Tests Comparing Test Scores Between Those Who Took the Practice Tests Plus the Actual Tests and Those Who Took Only the Actual Tests

	Took practice tests[a]		Did not take practice tests[b]			
Actual test type	*M*	*SD*	*M*	*SD*	*t*	*d*
Combined	102.96	14.33	98.47	17.53	18.77**	.27
Knowledge	50.27	8.78	48.44	9.83	12.90**	.19
English skills	52.69	7.59	50.03	9.68	20.65**	.29

[a] $N = 6{,}169$. [b] $N = 12{,}843$.
** $p < .01$.

Our research question asked whether within-subgroup gains in scores between the practice tests and the actual tests would be greater for minorities or for nonminorities. As Table 4 indicates, within-subgroup gains were greater for Blacks and Hispanics for combined score differences between the practice and actual tests ($d = .24$ difference in gains between Blacks and Whites and $d = .16$ between Hispanics and Whites). Within-subgroup differences in scores were also significant for the English skills tests ($d = .31$ between Blacks and Whites and $d = .18$ between Hispanics and Whites) but not significant for the knowledge tests ($d = .04$ between Blacks and Whites and $d = .05$ between Hispanics and Whites). To put these results in practical terms, we used hypothetical estimates of subgroup differences found in the previous research literature (Hough et al., 2001; Roth et al., 2017) and hypothetical passing rates. Assuming 1 *SD* as a mean-score difference between Blacks and Whites, equal numbers of candidates (100 Blacks and 100 Whites), and a passing rate of 50%, the practice test would result in the adverse impact ratio's increasing from .46 to .54 (17.39% increase). Performing the same analysis for Hispanics versus Whites assuming .5 *SD* as a mean-score difference, the practice test would result in the adverse impact ratio's increasing from .67 to .75 (11.94% increase). Although these estimates are hypothetical and not based on actual data from the organization, they illustrate the potential benefit in terms of adverse impact reduction.

Table 2
Means, Standard Deviations, and Intercorrelations

Variable	*N*	*M*	*SD*	1	2	3	4	5	6	7	8
1. Actual test: Combined	19,012	99.93	16.69	—							
2. Actual test: Knowledge	19,012	49.03	9.54	.90**	—						
3. Actual test: English skills	19,012	50.89	9.13	.89**	.60**	—					
4. Initial practice test: Combined	23,238	94.57	18.42	.77**	.65**	.71**	—				
5. Initial practice test: Knowledge	23,238	48.72	9.06	.68**	.65**	.53**	.87**	—			
6. Initial practice test: English skills	23,238	45.85	11.48	.67**	.50**	.70**	.92**	.60**	—		
7. Score gains: Combined tests	6,169	1.75	10.05	.27**	.30**	.16**	−.41**	−.32**	−.40**	—	
8. Score gains: Knowledge test	6,169	−1.25	6.98	.36**	.52**	.07**	−.16**	−.30**	.00	.74**	—
9. Score gains: English skills test	6,169	3.01	6.76	.03**	−.09**	.17**	−.45**	−.16**	−.59**	.72**	.07**

Note. Score gains are between practice tests and actual tests.
* $p < .05$. ** $p < .01$.

Table 4
Subgroup Differences in Score Gains Between Practice and Actual Tests

Tests taken	Blacks[a]		Hispanics[b]		Whites[c]			
	M_{diff}	*SD*	M_{diff}	*SD*	M_{diff}	*SD*	*t*	*d*
Actual tests (Combined)—Practice tests (Combined)	3.68	13.20			1.34	9.44	3.92**	.24
Actual knowledge test—Practice knowledge test	−1.07	8.50			−1.33	6.71	.66	.04
Actual English skills test—Practice English skills test	4.75	9.18			2.67	6.30	5.03**	.31
Actual tests (Combined)—Practice tests (Combined)			2.83	10.74	1.34	9.44	3.22**	.16
Actual knowledge test—Practice knowledge test			−1.00	7.42	−1.33	6.71	1.02	.05
Actual English skills test—Practice English skills test			3.83	7.30	2.67	6.30	3.71**	.18

[a] $N = 522$. [b] $N = 601$. [c] $N = 4{,}286$.
** $p < .01$.

Hypothesis 2 predicted that those who scored higher on the practice tests would be more likely to apply. This hypothesis was supported for scores on both practice tests combined ($r_{pb} = .22$, $p < .01$), as well as for the individual practice knowledge ($r_{pb} = .19$, $p < .01$) and English skills ($r_{pb} = .20$, $p < .01$) tests. To gain a practical interpretation of these effects, we analyzed them using logistic regression. Results indicated that a 1-point increase was associated with a 3% increase in likelihood of applying for the combined practice test, 5% for the job knowledge test, and 5% for the English skills test. In addition, we examined how practice test scores related to decisions to apply by percentile. The results for this analysis show that increases in practice test scores by decile are systematically associated with increases in likelihood of applying from 12% likelihood of applying for those in the bottom decile to 43% for the top decile (see Table 5).

Hypothesis 3 predicted that the gains between the practice and actual tests would be greater than the gains associated with retesting. As Table 6 indicates, this hypothesis was not supported. There was not a significant difference between practice and actual tests in combined scores ($d = -.05$). Upon examining the gains between practice testers and retesters by type of test, the results indicated that this hypothesis was supported for score gains on the English skills test ($d = .25$). However, the opposite trend was observed for the knowledge test, where retesting was more beneficial ($d = -.32$). In other words, practice testing in this context is more beneficial for scores on the English skills test, whereas retesting is more beneficial for scores on the knowledge test. In practical terms, assuming an average of 15,000 candidates per year and a 50% passing rate, the practice test enables approximately 375 additional applicants to pass the English skills assessment, whereas retesting enables approximately 450 additional applicants to pass the knowledge assessment.

Table 5
Percentage of Practice Testers Applying by Percentile Score on Practice Tests

Practice test score percentile	*N*	No. applying	% applying
10	2,323	272	11.71
20	2,297	352	15.32
30	2,296	390	16.99
40	2,421	446	18.84
50	2,219	494	22.26
60	2,382	635	26.66
70	2,316	762	32.90
80	2,341	843	36.01
90	2,324	962	41.39
100	2,319	1,003	43.25

Note. $N = 23{,}238$. For this analysis, the total practice test score was used. This was computed as the sum of the knowledge and English skills tests.

Supplemental Analyses

In addition to the hypothesis tests discussed, we conducted several additional analyses to consider potential alternative explanations for our findings.

Are gains observed between the practice and actual tests due to lack of motivation on the practice test? To examine this issue, we compared the ranges and distributions of the practice test scores to the actual test scores. If there were differences in test taker motivation across the two groups, then the ranges in scores would be greater for the practice test and the scores for the actual test would be more negatively skewed and leptokurtic than those for the practice test. Results regarding the ranges indicated that, when summing the tests, the range of scores for the practice tests and actual tests were similar (13.32 to 129.10 and 21.70 to 133.01, respectively). Results were similar at the subtest level (9.61 to 66.95 for the practice knowledge test and 11.82 to 70.24 for the actual knowledge test; 2.15 to 62.14 for the practice English skills test and 1.79 to 63.37 for actual English skills test). Results regarding the score distributions indicated they were similar for each group in terms of their skew (−1.49 vs. −1.22 for practice

Table 6
T Tests Comparing Test Score Gains Between Those Who Took the Practice Tests and Those Who Retested

Score gains	Took practice tests[a]		Retested[b]			
	M	*SD*	*M*	*SD*	*t*	*d*
Combined tests	1.75	10.05	2.29	8.15	1.34	−.05
Knowledge test	−1.25	6.98	.93	5.60	7.84**	−.32
English skills test	3.01	6.76	1.37	5.45	6.05**	.25

[a] $N = 6{,}169$. [b] $N = 451$.
** $p < .01$.

tests combined and actual tests combined, −1.01 vs. −.64 for knowledge tests, and −1.76 vs. −1.73 for English skills tests) and kurtosis (3.12 vs. 2.34, 1.63 vs. .44, and 4.19 vs. 4.58, respectively). Thus, the results suggest that the lower stakes setting of the practice test does not reduce practice test takers' motivation to perform on the practice test, based on the test score distributions.

Is taking the practice test an indicator of just preparation and/or higher potential applicant quality? This question is important because it addresses whether the act of taking the practice test improves individuals' human capital or just identifies individuals who are generally more prepared and/or have higher levels of human capital already. On the one hand, some individuals are more prone to practice test taking and preparation in general. On the other hand, as noted earlier, the purpose of this practice test in the present context is explicitly to encourage weaker or "borderline" candidates to use it so they can determine their odds of being hired and reduce their potential disappointment.

We examined this question in two ways. First, we compared the means on the practice tests to the means for these tests when they were previously used as actual tests. Results indicated that the means for practice test takers were lower than the means when the practice tests were used previously as actual tests ($M = 94.57$, 48.72, and 45.85, for practice tests combined, practice knowledge, and practice English skills, respectively, compared to $M = 100$, 50, and 50 when used as actual tests, because these assessments were *T*-scored). Second, we compared the average biodata scores for applicants who took the practice test to scores for those who did not take the practice tests and found that those who did not take the practice tests had higher biodata scores ($M = 49.77$ vs. 50.79, $SD = 9.68$ vs. 9.94, $N = 6{,}169$ vs. 12,843, respectively), $t(.10) = 6.63$, $p < .01$. Altogether, results from these analyses suggest that practice testers are not more motivated and of higher quality as a group when compared to other applicants and, in fact, had somewhat lower scores.

Discussion

The inadvertent attraction of a large quantity of unqualified applicants remains a lasting problem for organizations recruiting for professional occupations (Dineen & Williamson, 2012). This quality–quantity dilemma results in less-than-optimal outcomes for both organizations and job seekers during personnel selection. Resolving this dilemma requires recognizing the interconnectedness of the goals between (a) recruitment and personnel selection and (b) organizations and job seekers. This study explores one potential partial solution. The purpose of this study was to develop a more holistic framework that attempts to identify and explain the advantages realized by both an organization and its (potential) applicants throughout recruitment and personnel selection that result from the use of practice tests and then examine its predictions and an associated research question using a large sample of candidates in an operational selection context. Results indicated that candidates who took the practice test scored higher on the actual test and that candidates who scored higher on the practice test were more likely to apply. These outcomes, in combination, should result in a slightly smaller, but more highly qualified, applicant pool. Results also indicated that practice tests are associated with greater score gains for minorities and that the gains from practice tests, overall, are comparable to those associated with retesting.

Theoretical Implications

Conceptualizing practice tests as a means of both signaling information regarding an organization's selection procedures and creating an opportunity for short-term human capital investment has broad theoretical implications. First, the present study extends recent work suggesting that organizations may be capable of enhancing the quality of their applicants in the long term (Campion et al., 2017) by identifying a practice that is capable of enhancing the quality of their applicant pools in the short term. Future research should begin to examine whether and how such recruitment practices aimed at enabling long- versus short-term human capital investments can be integrated into one overall recruitment strategy.

Second, our framework extends prior research on personnel selection by highlighting the need to think more broadly about how to mitigate the adverse impact–validity dilemma. Prior research on this topic has typically focused on how to improve personnel selection practices (Campion et al., 2016; Ployhart & Holtz, 2008; Sackett & Ellingson, 1997; Sackett & Roth, 1996). However, our study suggests that it may also be useful to look "upstream" of the selection process to recruitment practices and to consider how they might be used to reduce subsequent subgroup differences in selection outcomes. For example, in this study we not only demonstrated that Blacks and Hispanics tended to realize greater gains from practice tests but also used previous estimates of subgroup differences from the literature to provide preliminary estimates of reduced adverse impact. Future research should seek to replicate this finding in other settings and with other selection procedures.

Third, we extend prior research on recruitment that focuses on the use of realistic information to influence (potential) applicant decisions. We demonstrated that, complementing research demonstrating that realistic job previews can reduce turnover, information acquired regarding an organization's selection procedures (i.e., through taking practice tests) influences immediate application decisions. Individuals scoring higher on the practice test were more likely to apply. Future research should examine whether practice tests make candidates more resistant to dropping out of the hiring process if they are unsuccessful because they provide a way for candidates to measure their improvement and determine whether to apply again.

Fourth, we compared practice testing and retesting in terms of the score gains realized by applicants. Signaling theory proposes that signals that are more costly to signal senders are more valuable to signal receivers (Connelly et al., 2011; Spence, 1973). Further, researchers applying signaling theory to understand staffing have suggested that organizations should not signal information regarding their selection procedures–criteria, because it may offer certain applicants an unfair advantage or, worse, may damage the integrity of the selection process (Bangerter et al., 2012). By showing that score gains from practice tests are comparable to score gains from retesting, despite their large differences in costs to the organization, we therefore highlight a potential boundary condition to the application of signaling theory in a staffing context when it comes to signal cost. Specifically, we provide evidence that there exist contexts in which less costly signals may be

just as valuable and identify a condition under which revealing information regarding assessment criteria may be helpful to the organization. In light of recent theory suggesting that applicants and organizations are involved in adaptive signaling games where each party continually escalates its competitive behavior (Bangerter et al., 2012), future research should examine whether practice testing and retesting differ in terms of which might lead to negative outcomes. For example, it is possible that failing the actual tests has a more lasting negative effect on applicants, leading to less favorable reactions and employer image perceptions, which, in turn would reduce the value of this signal further for the organization.

In addition, the present study yielded two somewhat unexpected findings. First, we found that score gains between the practice and actual tests occurred for the English skills test but not for the knowledge test. Second, our results indicated that score gains on the knowledge test did not occur between practice and actual knowledge tests, but they did occur for retesters. There are several possible explanations for these findings. For example, this may have occurred because applicants who took the practice test were slightly weaker candidates, as the results from one of the supplemental analyses indicated. Therefore, when practice test takers subsequently take actual tests, they do not improve in a normative sense compared to other candidates but in fact appear to perform slightly less well. On the other hand, it may be that meaningful human capital gains cannot occur in a short time period where knowledge needed to pass assessments not only covers a large number of knowledge dimensions (20 in this case) but also must be sought out and learned only after one extrapolates them from the test. Thus, gains on the knowledge test may require a longer period of time, which is consistent with the results found for retesters. Note that retesters can retest only once a year at this organization, whereas the practice test takers can take the practice test anytime, including immediately prior to taking the actual test (although the mean was 70.84 days). The latter explanation may also explain why larger gains were found for English skills for the practice test compared to retesting. In this case, the applicants only needed to refamiliarize themselves with a much smaller number of "rules" with which they were likely already acquainted. They also would likely benefit from a shorter amount of time between practice test and application because this skill may fade as time passes. Future research should explore the use of a parallel follow-up practice test candidates can take, which requires that a longer period of time has passed (e.g., 6 months). This would provide a useful tool for candidates seeking to improve their scores to make themselves more competitive because they may be capable of better gauging their actual increases in human capital.

Practical Implications

Firms face many challenges when recruiting for professional occupations today. One of the most critical is that they are often inundated with a large volume of underqualified applicants. This results in numerous problems, ranging from a general inability to amass enough talent, to increased pressure and costs placed on selection systems to screen out the many unqualified candidates, to increased legal exposure from the large number of rejected candidates. The present study introduces the use of practice tests for the purposes of recruitment as a potential way to combat these issues and offers several key implications for organizations. The primary practical implication is that organizations should begin offering practice tests during recruitment because they provide one way to improve the quality of applicants in the short term. Practice tests provide an opportunity for applicants nearing (or on) the job market to make short-term investments toward improving their human capital. They also provide a way for potential applicants to preview–learn about the organization's selection process and assess their chances of doing well enough to receive a job offer. Moreover, practice testing may reduce adverse impact for minorities because offering practice tests may provide a mechanism through which minority applicants can familiarize themselves with the selection process and increase their scores, as they did in this study. Finally, practice tests may be a lower cost alternative to allowing retesting but provide applicants with opportunities to achieve similar score gains overall.

The experiences of the organization studied here may be useful to other organizations that want to use practice tests. For example, in this organization, we found that when providing the option of taking practice tests; making them anonymous, available online, and available to everyone; and encouraging individuals to take them before applying yielded higher quality candidates. Practice tests not only required less effort and time on the part of the potential applicant than did applying directly but also ensured that a safe, low-stress environment was maintained wherein learning and development might be enhanced to a greater degree. Moreover, a key advantage for this organization was that normative data were available, which enabled it to provide more accurate estimates to potential applicants regarding their likelihood of success.

Limitations and Associated Future Research Directions

The present study had a number of strengths but was not without limitations. First, the framework developed in this study may be dependent on the level of rigor associated with the organization's hiring standards as well as the characteristics of the candidates themselves. The context examined in this study is somewhat unique in that the organization's hiring standards are extremely rigorous and candidates are highly motivated, actively seeking information and advice on how to get hired. Therefore, future research should attempt to replicate these findings in other organizations.

Second, although we offered initial theoretical explanations for the effects of practice tests, we did not directly test some of the constructs, particularly those related to our findings regarding racial−ethnic subgroup differences in score gains. We do not know the extent to which different mechanisms (e.g., enhanced test-taking attitudes, enhanced test-content familiarity, enhanced test-taking strategies) were responsible for the effects found. Future research should extend our findings by exploring how practice tests influence these constructs. In so doing, future research would complement the present study by providing more insight regarding the value of practice tests from the perspective of applicants.

Third, we were unable to access performance data on applicants who ultimately obtained jobs within this occupation. Therefore, we could not examine the effects of practice tests on criterion-related validity or on the quality of new hires. Future research should

examine these issues because prior research has demonstrated that retesting can enhance criterion-related validity of tests (Van Iddekinge et al., 2011), and it is important to understand whether practice tests improve the quality of new hires and to what degree.

Fourth, this research was limited to knowledge and skills tests. The findings would likely be applicable to some other types of tests and methods that measure KSAs, such as situational judgment tests, and can be administered online. The logic of providing feedback can be extended to tests measuring other characteristics (e.g., personality) but not the logic of providing recommendations for improvement, because those attributes are thought to be stable and immutable. Moreover, there is some risk that the feedback would coach candidates into faking their responses (Levashina et al., 2009). Practice tests may also not be as applicable to methods that require a hiring official to administer, such as an interview or assessment exercise. Interview questions are also highly susceptible to faking (Levashina & Campion, 2007) and perhaps more so if the questions are revealed in advance. As such, the generalizability of the findings to other types of tests and methods is a fertile area for future research.

Finally, because practice test data and actual test data were matched using e-mail addresses, it is possible that some number of individuals who took the practice test also took the actual test but were unmatched in the data due to the use of different e-mail addresses for the practice and actual tests. However, there is little reason to believe that using a different e-mail address would be related to test scores in any manner. For example, it was explained to practice test takers that their scores were confidential and kept anonymous and that the practice test was for their benefit and not part of the selection process, so test takers had little incentive to further anonymize their practice test scores. Moreover, the average time between taking the practice test and the actual test was relatively short, so e-mails were unlikely to change.

In conclusion, we proposed that the use of practice tests both reduces information asymmetry regarding an organization's selection procedures and serves as a short-term human capital investment option for potential applicants. Using this general framework, we showed that candidates who took advantage of practice tests during recruitment performed better during assessment, that minorities tended to realize greater gains between the practice and actual assessments, that practice test scores were associated with application decisions of potential applicants, and that practice tests may provide a low-cost alternative to retesting for organizations. We hope that future research will advance these initial findings by refining our theoretical arguments and measures.

References

Anastasi, A. (1981). Coaching, test sophistication, and developed abilities. *American Psychologist, 36,* 1086–1093. http://dx.doi.org/10.1037/0003-066X.36.10.1086

Ariely, D., Gneezy, U., Loewenstein, G., & Mazar, N. (2009). Large stakes and big mistakes. *Review of Economic Studies, 76,* 451–469. http://dx.doi.org/10.1111/j.1467-937X.2009.00534.x

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General, 130,* 224–237. http://dx.doi.org/10.1037/0096-3445.130.2.224

Avery, D. R., Hernandez, M., & Hebl, M. R. (2004). Who's watching the race? Racial salience in recruitment advertising. *Journal of Applied Social Psychology, 34,* 146–161. http://dx.doi.org/10.1111/j.1559-1816.2004.tb02541.x

Bangerter, A., Roulin, N., & König, C. J. (2012). Personnel selection as a signaling game. *Journal of Applied Psychology, 97,* 719–738. http://dx.doi.org/10.1037/a0026078

Barber, A. E. (1998). *Recruiting employees: Individual and organizational perspectives* (Foundations for Organizational Science, Book 8). http://dx.doi.org/10.4135/9781452243351

Bargh, J. A., Lombardi, W. J., & Higgins, E. T. (1988). Automaticity of chronically accessible constructs in person × situation effects on person perception: It's just a matter of time. *Journal of Personality and Social Psychology, 55,* 599–605. http://dx.doi.org/10.1037/0022-3514.55.4.599

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy, 70,* 9–49. http://dx.doi.org/10.1086/258724

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis with special reference to education*. Chicago, IL: University of Chicago Press.

Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological Science, 16,* 101–105. http://dx.doi.org/10.1111/j.0956-7976.2005.00789.x

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review, 89,* 1–47. http://dx.doi.org/10.1037/0033-295X.89.1.1

Breaugh, J. A., & Starke, M. (2000). Research on employee recruitment: So many studies, so many remaining questions. *Journal of Management, 26,* 405–434. http://dx.doi.org/10.1177/014920630002600303

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64,* 123–152. http://dx.doi.org/10.1037/h0043805

Campion, M. C. (2015, August). *Minorities versus nonminorities: Implications of subgroup signaling markets for targeted recruitment*. Paper presented at the annual conference of the Academy of Management, Vancouver, British Columbia, Canada.

Campion, M. C., Campion, M. A., Campion, E. D., & Reider, M. H. (2016). Initial investigation into computer scoring of candidate essays for personnel selection. *Journal of Applied Psychology, 101,* 958–975. http://dx.doi.org/10.1037/apl0000108

Campion, M. C., Ployhart, R. E., & Campion, M. A. (2017). Using recruitment source timing and diagnosticity to enhance applicants' occupation-specific human capital. *Journal of Applied Psychology, 102,* 764–781. http://dx.doi.org/10.1037/apl0000190

Cappelli, P. (2001). Making the most of on-line recruiting. *Harvard Business Review, 79,* 139–146, 166.

Chan, D. (1997). Racial subgroup differences in predictive validity perceptions on personality and cognitive ability tests. *Journal of Applied Psychology, 82,* 311–320. http://dx.doi.org/10.1037/0021-9010.82.2.311

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management, 37,* 39–67. http://dx.doi.org/10.1177/0149206310388419

Dineen, B. R., & Noe, R. A. (2009). Effects of customization on application decisions and applicant pool characteristics in a web-based recruitment context. *Journal of Applied Psychology, 94,* 224–234. http://dx.doi.org/10.1037/a0012832

Dineen, B. R., & Williamson, I. O. (2012). Screening-oriented recruitment messages: Antecedents and relationships with applicant pool quality. *Human Resource Management, 51,* 343–360. http://dx.doi.org/10.1002/hrm.21476

Feldman, J. M., & Lynch, J. G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology, 73,* 421–435. http://dx.doi.org/10.1037/0021-9010.73.3.421

Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology, 57,* 639–683. http://dx.doi.org/10.1111/j.1744-6570.2004.00003.x

Hausknecht, J. P., Halpert, J. A., Di Paolo, N. T., & Moriarty Gerrard, M. O. (2007). Retesting in selection: A meta-analysis of coaching and practice effects for tests of cognitive ability. *Journal of Applied Psychology, 92,* 373–385. http://dx.doi.org/10.1037/0021-9010.92.2.373

Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment, 9,* 152–194. http://dx.doi.org/10.1111/1468-2389.00171

Kulik, J. A., Kulik, C.-L. C., & Bangert, R. L. (1984). Effects of practice on aptitude and achievement test scores. *American Educational Research Journal, 21,* 435–447. http://dx.doi.org/10.3102/00028312021002435

Levashina, J., & Campion, M. A. (2007). Measuring faking in the employment interview: Development and validation of an interview faking behavior scale. *Journal of Applied Psychology, 92,* 1638–1656. http://dx.doi.org/10.1037/0021-9010.92.6.1638

Levashina, J., Morgeson, F. P., & Campion, M. A. (2009). They don't do it often, but they do it well: Exploring the relationship between applicant mental abilities and faking on biodata measures. *International Journal of Selection and Assessment, 17,* 271–281. http://dx.doi.org/10.1111/j.1468-2389.2009.00469.x

Levashina, J., Morgeson, F. P., & Campion, M. A. (2012). Tell me some more: Exploring how verbal ability and item verifiability influence responses to biodata questions in a high-stakes selection context. *Personnel Psychology, 65,* 359–383. http://dx.doi.org/10.1111/j.1744-6570.2012.01245.x

Lievens, F., Buyse, T., & Sackett, P. R. (2005). Retest effects in operational selection settings: Development and test of a framework. *Personnel Psychology, 58,* 981–1007. http://dx.doi.org/10.1111/j.1744-6570.2005.00713.x

McCarthy, J. M., Van Iddekinge, C. H., & Campion, M. A. (2010). Are highly structured job interviews resistant to demographic similarity effects? *Personnel Psychology, 63,* 325–359. http://dx.doi.org/10.1111/j.1744-6570.2010.01172.x

Messick, S., & Jungeblut, A. (1981). Time and method in coaching for the SAT. *Psychological Bulletin, 89,* 191–216. http://dx.doi.org/10.1037/0033-2909.89.2.191

Newman, D. A., & Lyon, J. S. (2009). Recruitment efforts to reduce adverse impact: Targeted recruiting for personality, cognitive ability, and diversity. *Journal of Applied Psychology, 94,* 298–317. http://dx.doi.org/10.1037/a0013472

Nguyen, H. H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology, 93,* 1314–1334. http://dx.doi.org/10.1037/a0012702

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (McGraw-Hill Series in Psychology, Vol. 3). New York, NY: McGraw-Hill.

Phillips, J. M. (1998). Effects of realistic job previews on multiple organizational outcomes: A meta-analysis. *Academy of Management Journal, 41,* 673–690.

Ployhart, R. E., & Holtz, B. C. (2008). The diversity–validity dilemma: Strategies for reducing racioethnic and sex subgroup differences and adverse impact in selection. *Personnel Psychology, 61,* 153–172. http://dx.doi.org/10.1111/j.1744-6570.2008.00109.x

Powers, D. E. (1986). Relations of test item characteristics to test preparation/test practice effects: A quantitative summary. *Psychological Bulletin, 100,* 67–77. http://dx.doi.org/10.1037/0033-2909.100.1.67

Rawson, K. A., & Dunlosky, J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology: General, 140,* 283–302. http://dx.doi.org/10.1037/a0023956

Ree, M. J., & Earles, J. A. (1991). Predicting training success: Not much more than g. *Personnel Psychology, 44,* 321–332. http://dx.doi.org/10.1111/j.1744-6570.1991.tb00961.x

Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17,* 249–255. http://dx.doi.org/10.1111/j.1467-9280.2006.01693.x

Roth, P. L., Van Iddekinge, C. H., DeOrtentiis, P. S., Hackney, K. J., Zhang, L., & Buster, M. A. (2017). Hispanic and Asian performance on selection procedures: A narrative and meta-analytic review of 12 common predictors. *Journal of Applied Psychology, 102,* 1178–1202. http://dx.doi.org/10.1037/apl0000195

Ryan, A. M. (2001). Explaining the Black-White test score gap: The role of test perceptions. *Human Performance, 14,* 45–75. http://dx.doi.org/10.1207/S15327043HUP1401_04

Rynes, S. L., Bretz, R. D., Jr., & Gerhart, B. (1991). The importance of recruitment in job choice: A different way of looking. *Personnel Psychology, 44,* 487–521. http://dx.doi.org/10.1111/j.1744-6570.1991.tb02402.x

Sackett, P. R., Burris, L. R., & Ryan, A. M. (1989). Coaching and practice effects in personnel selection. In C. L. Cooper & I. Robertson (Eds.), *International review of industrial and organizational psychology* (pp. 145–183). New York, NY: Wiley.

Sackett, P. R., & Ellingson, J. E. (1997). The effects of forming multipredictor composites on group differences and adverse impact. *Personnel Psychology, 50,* 707–721. http://dx.doi.org/10.1111/j.1744-6570.1997.tb00711.x

Sackett, P. R., & Roth, L. (1996). Multi-stage selection strategies: A Monte Carlo investigation of effects on performance and minority hiring. *Personnel Psychology, 49,* 549–572. http://dx.doi.org/10.1111/j.1744-6570.1996.tb01584.x

Schleicher, D. J., Van Iddekinge, C. H., Morgeson, F. P., & Campion, M. A. (2010). If at first you don't succeed, try, try again: Understanding race, age, and gender differences in retesting score improvement. *Journal of Applied Psychology, 95,* 603–617. http://dx.doi.org/10.1037/a0018920

Schleicher, D. J., Venkataramani, V., Morgeson, F. P., & Campion, M. A. (2006). So you didn't get the job . . . now what do you think? Examining opportunity-to-perform fairness perceptions. *Personnel Psychology, 59,* 559–590. http://dx.doi.org/10.1111/j.1744-6570.2006.00047.x

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics, 87,* 355–374. http://dx.doi.org/10.2307/1882010

Spence, M. (2002). Signaling in retrospect and the informational structure of markets. *American Economic Review, 92,* 434–459. http://dx.doi.org/10.1257/00028280260136200

Stiglitz, J. E. (2000). The contributions of the economics of information to twentieth century economics. *Quarterly Journal of Economics, 115,* 1441–1478. http://dx.doi.org/10.1162/003355300555015

Strober, M. H. (1990). Human capital theory: Implications for HR managers. *Industrial Relations, 29,* 214–239. http://dx.doi.org/10.1111/j.1468-232X.1990.tb00752.x

Tomlinson, M. (2007). Graduate employability and student attitudes and orientations to the labour market. *Journal of Education and Work, 20,* 285–304. http://dx.doi.org/10.1080/13639080701650164

Tomlinson, M. (2008). Students' perceptions of the role of higher education credentials for graduate work and employability. *British Journal of Sociology of Education, 29,* 49–61. http://dx.doi.org/10.1080/01425690701737457

Uggerslev, K. L., Fassina, N. E., & Kraichy, D. (2012). Recruiting through the stages: A meta-analytic test of predictors of applicant attraction at

different stages of the recruiting process. *Personnel Psychology, 65,* 597–660. http://dx.doi.org/10.1111/j.1744-6570.2012.01254.x

Uniform guidelines on employee selection procedures. (1978). *Federal Register, 43,* 38290–38315.

U.S. Department of Labor, Office of Federal Contract Compliance Programs. (2014). *Federal contract compliance manual*. Retrieved from https://www.dol.gov/ofccp/regs/compliance/fccm/FCCM_FINAL_508c.pdf

Van Hoye, G., & Lievens, F. (2009). Tapping the grapevine: A closer look at word-of-mouth as a recruitment source. *Journal of Applied Psychology, 94,* 341–352. http://dx.doi.org/10.1037/a0014066

Van Iddekinge, C. H., Morgeson, F. P., Schleicher, D. J., & Campion, M. A. (2011). Can I retake it? Exploring subgroup differences and criterion-related validity in promotion retesting. *Journal of Applied Psychology, 96,* 941–955. http://dx.doi.org/10.1037/a0023562

Zigerell, L. J. (2017). Potential publication bias in the stereotype threat literature: Comment on Nguyen and Ryan (2008). *Journal of Applied Psychology, 102,* 1159–1168. http://dx.doi.org/10.1037/apl0000188

Zottoli, M. A., & Wanous, J. P. (2000). Recruitment source research: Current status and future directions. *Human Resource Management Review, 10,* 353–382. http://dx.doi.org/10.1016/S1053-4822(00)00032-2

Received April 10, 2017
Revision received January 19, 2019
Accepted January 25, 2019 ■