9-2013

Do candidate reactions relate to job performance or affect criterion-related validity? A multistudy investigation of relations among reactions, selection test scores, and job performance

Julie MCCARTHY
University of Toronto

Chad H. VAN IDDEKINGE
Florida State University

Filip LIEVENS
Singapore Management University, filiplievens@smu.edu.sg

Mavis Mei-Chuan KUNG
Select International

Evan F. SINAR
Development Dimensions International

See next page for additional authors

DOI: https://doi.org/10.1037/a0034089

Do candidate reactions relate to job performance or affect criterion-related validity? A multistudy investigation of relations among reactions, selection test scores, and job performance

McCarthy, Julie M. Department of Management, University of Toronto–Scarborough, Toronto, ON, Canada, julie.mccarthy@rotman.utoronto.ca

Van Iddekinge, Chad H. College of Business, Department of Management, Florida State University, FL

Lievens, Filip, ORCID 0000-0002-9487-5187. Department of Personnel, Work, and Organizational Psychology, Ghent University, Ghent, Belgium

Kung, Mei-Chuan. Select International, Pittsburgh, PA

Sinar, Evan F. Development Dimensions International, Bridgeville, PA

Campion, Michael A. Krannert School of Management, Purdue University, IN

Published in Journal of Applied Psychology, Sep 2013, 98 (5), 701-719.

https://doi.org/10.1037/a0034089

Accepted version

Abstract

Considerable evidence suggests that how candidates react to selection procedures can affect their test performance and their attitudes toward the hiring organization (e.g., recommending the firm to others). However, very few studies of candidate reactions have examined one of the outcomes organizations care most about: job performance. We attempt to address this gap by developing and testing a conceptual framework that delineates whether and how candidate reactions might influence job performance. We accomplish this objective using data from 4 studies (total N = 6,480), 6 selection procedures (personality tests, job knowledge tests, cognitive ability tests, work samples, situational judgment tests, and a selection inventory), 5 key candidate reactions (anxiety, motivation, belief in tests, self-efficacy, and procedural justice), 2 contexts (industry and education), 3 continents (North America, South America, and Europe), 2 study designs (predictive and concurrent), and 4 occupational areas (medical, sales, customer service, and technological). Consistent with previous research, candidate reactions were related to test scores, and test scores were related to job performance. Further, there was some evidence that reactions affected performance indirectly through their influence on test scores. Finally, in no cases did candidate reactions affect the prediction of job performance by increasing or decreasing the criterion-related validity of test scores. Implications of these findings and avenues for future research are discussed.

Acknowledgement: A portion of this research was presented at the annual conference of the Society for Industrial and Organizational Psychology, Atlanta, Georgia, April 2010. This research was supported, in part, by Research Grant 410-2011-0313 from the Social Sciences and Humanities Research Council of Canada, awarded to Julie M. McCarthy. We would like to thank Herman Aguinis, Talya Bauer, and Donald Truxillo for their helpful comments on a preliminary version of this paper.

Recent years have witnessed an emerging body of research on candidate reactions to selection processes. These reactions reflect how candidates perceive and respond to selection tools (e.g., standardized tests, interviews, situational judgment tests) and include feelings of anxiety, levels of motivation, belief in tests, levels of self-efficacy, and perceptions of fairness and justice (Ryan & Ployhart, 2000). Consideration of candidate reactions is crucial, as reactions have been found to have notable implications for scores on selection procedures, as well as for candidates' subsequent attitudes and intentions. For example, candidates with positive reactions are more likely to be motivated during the selection process, perceive the organization as an attractive place to work, and recommend the organization to others (Hausknecht, Day, & Thomas, 2004).

However, a critical gap in the research on candidate reactions is the lack of attention to the outcome that many organizations care most about: job performance. This is somewhat surprising, as it is widely acknowledged that examining the role of candidate reactions with respect to job performance is important (e.g., Chan & Schmitt, 2004; Hülsheger & Anderson, 2009; McCarthy, Hrabluik, & Jelley, 2009; Reeve, Heggestad, & Lievens, 2009; Truxillo & Bauer, 2010). In fact, job performance is present in most theoretical models of candidate reactions (e.g., Gilliland, 1993; Hausknecht et al., 2004; Ryan & Ployhart, 2000). For example, Gilliland's seminal model proposes that candidate perceptions of selection process fairness will directly influence their attitudes (e.g., toward the organization), intentions (e.g., to recommend the selection process to others), and behaviors (e.g., test performance, job performance). However, to our knowledge, only one study (Gilliland, 1994) has found evidence of a direct link between candidate reactions and job performance.

It also has been suggested that candidate reactions may affect relations between test scores and job performance, or the criterion-related validity of selection procedures (e.g., Arvey, Strickland, Drauden, & Martin, 1990; McCarthy et al., 2009; Reeve et al., 2009; Ryan & Ployhart, 2000; Schmit & Ryan, 1992). For example, negative reactions (i.e., high anxiety) may introduce extraneous variance into test scores and, in turn, reduce the extent to which the scores reflect candidates' true level of ability and/or the extent to which the scores predict future performance. As a result, negative reactions may unduly influence test performance and, for example, lead to the rejection of high-quality job applicants and/or the acceptance of low-quality applicants. Indeed, this proposition is often raised as an argument for why the study of candidate reactions is so important (e.g., Hülsheger & Anderson, 2009; Reeve et al., 2009; Smither, Reilly, Millsap, & Pearlman, 1993). Unfortunately, very little research has examined these possibilities, which has led to various calls for studies that examine the effects of applicant reactions on the relation between test scores and performance on the job (e.g., Chan & Schmitt, 2004; Morgeson & Ryan, 2009; Ryan & Ployhart, 2000; Sackett & Lievens, 2008).

The current investigation responds to these calls by examining whether and how candidate reactions are related to job performance. This study adds to the selection literature in four main ways. First, existing studies have typically focused on how reactions relate to test performance (Arvey et al., 1990; Chan, Schmitt, Jennings, Clause, & Delbridge, 1998; Smither et al., 1993) or to candidate attitudes, such as intentions to remain in the selection process or to accept the position (e.g., Bauer, Maertz, Dolen, & Campion, 1998; Macan, Avedon, Paese, & Smith, 1994; Truxillo, Bauer, Campion, & Paronto, 2002). Although these studies provide a useful starting point, they do not inform whether or how candidate reactionsinfluence how well selection procedures predict how employees perform on the job. We address this critical question by linking reactions to scores on selection tests and measures of job performance.

Second, we draw from theory and past research to develop a conceptual framework that delineates how candidate reactions may affect test performance, job performance, and the relationship between the two variables. Providing conceptual clarification of this issue is valuable, as there are various possibilities concerning how candidate reactions may (or may not) contribute to the prediction of criteria such as performance. We also outline various boundary conditions that may affect when and how reactions affect test performance, job performance, and the relationship between the two variables.

Third, investigations of candidate reactions have tended to focus on perceptions of justice (e.g., Gilliland, 1994; Truxillo et al., 2002) or anxiety (e.g., McCarthy & Goffin, 2004; Proost, Derous, Schreurs, Hagtvet, & De Witte, 2008). We expand this focus and integrate insights from several research streams (i.e., anxiety, motivation, belief

in tests, self-efficacy, and justice) to assess candidate reactions from each of these domains. This step is critical to understanding the full range of reactions to selection procedures and the extent to which certain reactions may be more or less likely to influence test performance, job performance, and criterion-related validity. Our examination of multiple reactions also responds to the calls for the integration of research on test attitudes and fairness (e.g., Ryan & Ployhart, 2000).

Fourth, we consider the effects of candidate reactions to many commonly used types of selection procedures that measure a wide range of constructs, including "sign-based" (personality tests, job knowledge tests, cognitive ability tests) and "sample-based" (work samples and situational judgment tests [SJTs]) selection techniques (Wernimont & Campbell, 1968). This is important because reactions have been shown to vary across selection techniques (e.g., Anderson & Witvliet, 2008; Hausknecht et al., 2004; Ryan & Ployhart, 2000), and consideration of candidate reactions to sample-based selection procedures is relatively sparse (for a discussion, see Bauer & Truxillo, 2006, and Ryan & Ployhart, 2000). Further, we examine the role of reactions in different contexts (industry and education), continents (North America, South America, and Europe), and occupational areas (medical school students, sales professionals, customer service representatives, and product technicians), as well as with different study designs (predictive and concurrent). This methodological diversity and multiple operationalizations provide a valuable opportunity to test whether and how test reactions influence job performance and serves to increase the generalizability of the findings.

Candidate Reactions and Job Performance: A Conceptual Framework

One goal of our work was to develop a conceptual framework that delineates the various ways candidate reactions could affect job performance and/or the prediction thereof. This framework integrates theory and research from several literatures, including research on candidate reactions, personnel selection, work attitudes, cognitive processing, and personality. Our framework suggests that candidate reactions could affect job performance in three ways: (a) reactions could have a direct effect on job performance, (b) reactions could have an indirect effect on job performance (i.e., through test performance), and/or (c) reactions could moderate the relation between test performance and job performance. We outline each of these possibilities in turn below, and a summary of our conceptual framework is presented in Table 1.

Table 1
Framework for Understanding Relations Between Candidate Reactions and Job Performance

Characteristic	Reactions are directly related to job performance	Reactions are indirectly related to job performance	Reactions influence the relation between test performance and job performance
Description	Candidate reactions have a dispositional component. Given that individual dispositions have been found to predict job performance, candidate reactions may also predict job performance.	Candidate reactions influence job performance via their influence on test performance. Test performance is enhanced when reactions are positive (and candidates can devote sufficient cognitive resources to the task) and is diminished when reactions are negative (and candidates are unable to devote sufficient cognitive resources to the task). In turn, test performance is related to performance on the job.	Negative candidate reactions (e.g., high anxiety, low motivation) lead to off-task processing and may introduce extraneous variance into test performance. Thus, test scores will not reflect candidates' true KSAOs, which, in turn, attenuates the relation between test scores and job performance. In contrast, positive candidate reactions (e.g., high levels of motivation, self-efficacy, and beliefs in tests) lead to on-task and self-regulatory processing. Thus, test scores will reflect candidates' true KSAOs, which, in turn, increases test validities.
Illustration	Test reactions ↓ ↓ ↓ Job performance	Test reactions ↓ Test performance ↓ Job performance	Test reactions → Test performance ↓ ↓ Job performance
Boundary conditions	This possibility is more likely when (a) the test reaction is dispositional (b) test reactions and job performance are conceptually matched	This possibility is more likely when (a) the test reaction is situational	This possibility is more likely when (a) the test reaction is situational (b) test reactions and job performance are not conceptually matched
Relevant theories	Dispositional test anxiety (Chamorro-Premuzic et al., 2008) Personality and test motivation (Salgado et al., 1996)	Processing efficiency theory (Eysenck & Calvo, 1992) Attentional control theory (Eysenck et al., 2007) Elaboration likelihood model (Petty & Cacioppo, 1986)	Affective events theory (Weiss & Cropanzano, 1996) Attitude–behavior relations (Ajzen & Fishbein, 1977) Job performance theory (Campbell et al., 1993)

Note. KSAOs = knowledge, skills, abilities, and other characteristics.

Candidate Reactions Are Directly Related to Job Performance

The first possibility is that candidate reactions have a direct effect on job performance. For example, applicants who experience test anxiety also may experience anxiety on the job (Ryan, 2001; Truxillo & Bauer, 2010). Thus, anxiety may negatively affect both test performance and job performance. Similarly, individuals with low levels of test-taking motivation, test self-efficacy, or perceived procedural justice may experience low levels of work-related motivation and self-efficacy and may be more sensitive to unfairness on the job (Arvey et al., 1990; Cohen-Charash & Spector, 2001; Gilliland, 1993). Finally, individuals who do not believe in tests also may distrust organizational assessments, policies, or procedures. In turn, these work-related attitudes may result in lower levels of job performance. For example, meta-analytic evidence suggests that anxiety is negatively related to performance on the job (Ford, Cerasoli, Higgins, & Decesare, 2011), whereas general self-efficacy (Judge & Bono, 2001) and perceptions of justice (Colquitt, Conlon, Wesson, Porter, & Ng, 2001) are positively related to performance. Thus, candidate reactions may be directly related to job performance because factors such as anxiety, motivation, belief in tests, self-efficacy, and justice perceptions are themselves related to performance. This leads to our first hypothesis:

Hypothesis 1: Candidate reactions will demonstrate a direct effect on job performance. In particular, test-taking motivation, belief in tests, self-efficacy, and perceptions of procedural justice will be positively related to job performance, whereas test-taking anxiety will be negatively related to job performance.

An important boundary condition of relations between candidate reactions and job performance may be the extent to which reactions are dispositional or situational in nature. The idea that candidate reactions have a dispositional component rests on the notion that reactions are related to stable individual differences in constructs such as personality (Chan & Schmitt, 2004; Gilliland, 1993; Schmit & Ryan, 1992). The reactions that are most likely to display dispositional tendencies are test-taking anxiety, motivation, and self-efficacy, as they focus on relatively stable cognitions and personality traits (Chan & Schmitt, 2004; Ryan & Ployhart, 2000; Schmit & Ryan, 1992). For example, test anxiety has been linked to neuroticism (e.g., Chamorro-Premuzic, Ahmetoglu, & Furnham, 2008), and test motivation has been linked to conscientiousness (e.g., Salgado, Remeseiro, & Iglesias, 1996). Given that conscientiousness, self-efficacy, and to a lesser extent neuroticism are related to performanceon the job (e.g., Barrick & Mount, 1991; Judge & Bono, 2001), these test-taking attitudes also may predict performance (Chan & Schmitt, 2004; Chan, Schmitt, DeShon, Clause, & Delbridge, 1997; Schmit & Ryan, 1992).

In contrast, belief in tests and procedural justice perceptions reflect situational-specific affective states that focus on applicant perceptions of the specific test procedures and decisions and have no obvious overlap with individual differences such as personality (Chan & Schmitt, 2004). According to affective events theory (Weiss & Cropanzano, 1996), affective states are transient, time-bound experiences. Given that affect fluctuates over time, significant relations between affective reactions to a discrete event (e.g., taking a selection test) and subsequent behaviors (e.g., performance on the job) are likely to be observed only when they occur within a short time frame. This idea is consistent with theory on attitude–behavior relations (e.g., Ajzen & Fishbein, 1977), which suggests that the correspondence between two variables (e.g., in terms of time) is an important determinant of the strength of their relation. Furthermore, models of job performance (e.g., Campbell, McCloy, Oppler, & Sager, 1993) treat knowledge, skills, abilities, and other characteristics (KSAOs) as the most direct determinants of performance and affective reactions to workplace events as more distal. Thus, according to these theoretical perspectives, belief in tests and procedural justice perceptions are not expected to be related to job performance because such reactions are situational specific and are not likely to affect how employees feel and behave once on the job.
Hypothesis 1a: Relations between candidate reactions and job performance will be stronger when candidate reactions are dispositional (i.e., anxiety, motivation, and self-efficacy) and weaker when they are situational (i.e., belief in tests and procedural justice perceptions).

An additional boundary condition that may influence relations between candidate reactions and job performance is the extent of conceptual similarity between reactions and performance. This is consistent with Ajzen and Fishbein's (1977) compatibility principle, which asserts that relations will be strongest when constructs are conceptually matched. For example, there is considerable evidence demonstrating that criterion-related validity is enhanced when predictors and criteria are conceptually aligned (e.g., Lievens, Buyse, & Sackett, 2005; Murphy & Shiarella, 1997; Schneider, Hough, & Dunnette, 1996; Tett, Jackson, & Rothstein, 1991). As an example, the validity of personality tests tends to be stronger when the traits are chosen based on a job analysis (Tett et al., 1991).

In a similar way, test reactions may have a stronger effect on job performance when the conceptual overlap between the reactions and performance is high. For example, test-taking anxiety may be more likely to affect job performance when the performance ratings measure adaptability or the ability to handle stress. Similarly, test-taking motivation may be more likely to affect job performance when the performanceratings focus on "will do" performance rather than on "can do" performance (Borman, White, Pulakos, & Oppler, 1991; Schmitt, Cortina, Ingerick, & Wiechmann, 2003). It is also likely that test-taking anxiety and motivation will have stronger relations with task-based job performance than with interpersonal-based job performance because, as described in detail in the next section, these reactions are related to how individuals allocate their attention, which can have significant effects on task performance (Eysenck, Derakshan, Santos, & Calvo, 2007). In contrast, perceptions of justice are likely to have stronger relations with interpersonal-based performance because high levels of procedural justice have been found to enhance social exchange relations and to motivate employees to reciprocate by engaging in interpersonal-based behaviors, such as organizational citizenship behaviors (Masterson, Lewis, Goldman, & Taylor, 2000; Posthuma & Campion, 2005).

Hypothesis 1b: Relations between candidate reactions and job performance will be stronger when candidate reactions and performance are conceptually matched and weaker when they are more distinct.

Candidate Reactions Are Indirectly Related to **Job Performance**

A second possibility is that candidate reactions are related to job performance by virtue of their relation with test performance. Considerable research has shown that people who react positively to selection procedures tend to receive higher scores on the procedures and vice versa. For example, candidates who experience high levels of test motivation and/or perceptions of procedural justice tend to achieve higher test scores than candidates with low levels of motivation or justice (e.g., Bauer et al., 2006; Hausknecht et al., 2004). In contrast, candidates who experience high levels of test anxiety tend to achieve lower test scores than candidates who experience low levels of test anxiety (e.g., Hausknecht et al., 2004; McCarthy et al., 2009).

Several cognitive processing theories provide a basis for the relationship between test reactions and test performance, including attentional control theory (Eysenck et al., 2007), processing efficiency theory (Eysenck & Calvo, 1992), interference theory (Sarason, Sarason, & Pierce 1990), and the elaboration likelihood model (Petty & Cacioppo, 1986). These theories share the prediction that anxiety and motivation influence test performancethrough the allocation of attentional resources, such that performance is enhanced when individuals are able to devote sufficient cognitive resources to the task and is reduced when individuals are unable to do so. Thus, individuals with high levels of motivation and self-efficacy, as well as those who possess a strong belief in tests, may be more likely to focus their attention on the task, resulting in higher levels of performance (Arvey et al., 1990; Bandura, 1982; Chan, Schmitt, Sacco, & DeShon, 1998). In contrast, high levels of anxiety are likely to lead to off-task and peripheral processing, which, in turn, may reduce test performance. Thus, theory and research provide a strong basis for a relationship between candidate reactions and test performance.

In turn, there is overwhelming evidence that selection test scores can predict performance on the job (Schmidt & Hunter, 1998). For example, cognitive ability tests (Hunter & Hunter, 1984; Schmitt, Gooding, Noe, & Kirsch, 1984), personality tests (Barrick & Mount, 1991), job knowledge tests (Dye, Reck, & McDaniel, 1993), work samples (Roth, Bobko, & McFarland, 2005), and SJTs (McDaniel, Hartman, Whetzel, & Grubb, 2007) have been found to demonstrate criterion-related validity for predicting performance.

Thus, if candidate reactions are correlated with test scores, and if test scores are correlated with job performance, then reactions may have an indirect effect on performance (via their influence on test scores). In line with recent thinking, the precise nature of this effect may depend on whether candidate reactions also have some direct effect on job performance. There is precedence in the literature for distinguishing between *mediated effects* and *indirect effects* (e.g., Mathieu & Taylor, 2006; Preacher & Hayes, 2004; Zhao, Lynch, & Chen, 2010). To be consistent with this distinction, we propose that a mediated effect will be present when candidate reactions exhibit a direct relation with job performance, as well as an indirect relation with job performance through test scores. In contrast, an indirect effect will be present when candidate reactions affect job performance only indirectly (through test scores) and do not have a direct effect on performance.

As we hypothesized (H1a), test reactions are more likely to relate to job performance when they capture job-relevant dispositions than when they capture more situational factors that are time-bound and less likely to carry over to the job. Thus, we predict that test performance will *mediate*relations between candidate reactions and job performance when the reactions are dispositional, because reactions are expected to be relevant to job performance. In contrast, when reactions are situational (and thus not expected to be relevant to job performance), we predict that reactions will have an *indirect effect* on job performance.

Hypothesis 2a: When candidate reactions are dispositional (i.e., anxiety, motivation, and self-efficacy), test performance will mediate relations between candidate reactions and job performance.

Hypothesis 2b: When candidate reactions are situational (i.e., belief in tests and procedural justice perceptions), they will exhibit an indirect effect on job performance through test performance.

Candidate Reactions Influence Relations Between Test Performance and **Job Performance**

The preceding hypotheses predict that candidate reactions will affect test performance and/or job performance. A third possibility is that reactions affect the relationship between test performance and job performance. One possibility that has been mentioned is that reactions function as a moderator of test–performance relations (Arvey et al., 1990; Ryan & Ployhart, 2000; Schmit & Ryan, 1992). That is, when reactions are negative (e.g., high levels of anxiety), the relation between test performance and job performance may be reduced, and when reactions are positive (e.g., high levels of test motivation), the relation between test performance and job performance may be enhanced.

The few previous studies that have examined whether candidate reactions influence criterion-related validity have primarily been conducted with student samples and have yielded mixed findings. In a simulated selection situation with college students, Schmit and Ryan (1992) found that positive reactions increased the validity of a cognitive ability test and decreased the validity of a personality test for the criterion of grade point average (GPA). Thorsteinson and Ryan (1997) found that students' perceptions of test fairness increased the validity of a cognitive ability test for GPA but had no effect on the validity of the personality test. Reeve and colleagues found that test anxiety decreased the predictive validity of cognitive ability tests in a Monte-Carlo simulation (Reeve et al., 2009), as well as in a student sample that considered grades as the criterion (Bonaccio, Reeve, & Winford, 2012). O'Neill, Goffin, and Gellatly (2010) found that applicants' test-taking motivation did not consistently moderate the validity of a personality test.

In spite of these mixed findings, there is theoretical reason to believe that applicant reactions are likely to affect criterion-related validity. The mechanism underlying this proposition is that negative reactions (e.g., high levels of anxiety) may introduce extraneous variance into test scores such that reactions reduce the extent to which the scores reflect candidates' true KSAOs (Hausknecht et al., 2004; Lievens et al., 2005). Attentional control theory (Eysenck et al., 2007), processing efficiency theory (Eysenck & Calvo, 1992), interference theory (Sarason et al., 1990), and the elaboration likelihood model (Petty & Cacioppo, 1986) provide the basis for this mechanism and predict that when candidates experience negative reactions to the test, they are unable to devote sufficient cognitive resources to the task. The result is lower levels of test performance. This, in turn, will attenuate test validities because the observed scores will be less reflective of candidates' actual KSAOs. In contrast, high levels of motivation, self-efficacy, and beliefs in tests are likely to lead to on-task and self-regulatory processing. In turn, the resulting test scores will be more likely to reflect the KSAOs that the test was designed to assess. This will have the effect of increasing test validities. We draw from these cognitive processing theories to predict the following:

Hypothesis 3: Candidate reactions will moderate relations between test scores and job performance, such that test–job performance relations will be weaker for candidates with negative reactions and stronger for candidates with positive reactions.

The boundary conditions noted earlier may account for the mixed findings in studies that have assessed whether candidate reactions influence criterion-related validity. In particular, dispositional-based reactions may be less likely than situational-based reactions to moderate validity. This is because dispositional reactions are more closely tied to job performance, and hence are less likely to introduce extraneous variance into selection test scores. In contrast, situational reactions are time-bound experiences that are not expected to be related to distal outcomes, such as job performance (Ajzen & Fishbein, 1977; Weiss & Cropanzano, 1996). In other words, situational-specific reactions that are evoked in testing situations (e.g., belief in tests, procedural justice perceptions) are unlikely to affect the day-to-day job performance of most employees. As a result, they are more likely to introduce extraneous variance into test scores.

Hypothesis 3a: Candidate reactions will have a stronger effect on the test–job performance relation when they are situational (i.e., belief in tests and procedural justice perceptions) than when they are dispositional (i.e., anxiety, motivation, and self-efficacy).

The degree of conceptual overlap between test reactions and job performance also may serve as a boundary condition. When this overlap is high, candidate reactions are relevant to performance on the job and are less likely to introduce extraneous variance into test scores.

Thus, the higher the conceptual similarity between test reactions and the criterion, the less likely reactions are to affect criterion-related validity.

Hypothesis 3b: Candidate reactions will have a weaker effect on the test–job performance relation when test-taking reactions and job performance are conceptually matched than when they are not conceptually matched.

Research Context

Our conceptual framework serves as a foundation from which different models and boundary conditions may be tested. We begin to test this framework using four separate field studies, each of which measured test performance, candidate reactions, and job performance. The first three studies examine job applicants in high-stakes selection testing contexts. Study 1 comprises applicants who completed a knowledge-based admission exam for medical school. It tests the extent to which four different reactions (anxiety, motivation, self-efficacy, and belief in tests) relate to scores on the knowledge test and to performance in medical school. Studies 2 and 3 extend this work to employment contexts. In Study 2, managerial-level job applicants for a large multinational organization are examined, and the extent to which perceptions of procedural justice relateto a multipredictor selection inventory and to ratings of job performance is assessed. Study 3 considers job applicants for sales positions in a large financial institution and tests the extent to which applicant perceptions of procedural justice relate to a personality test, a cognitive ability test, and job performance. Study 4 focuses on customer service representatives and product technicians in a large retail organization. The extent to which the reactions (anxiety, motivation, and procedural justice) of these incumbents relate to both sign-based (a personality test) and sample-based selection procedures (a work sample and a SJT), and job performance is assessed. In this last study, we included a "concurrent" research design to triangulate our findings and enhance multiple operationalizations. In particular, current employees may be more candid about their reactions to tests than job applicants. Thus, there may be more variance in test reactions, which, in turn, may provide a better opportunity for some of the predicted relationships to emerge.

Study 1

Method

Participants and procedure

The sample consisted of 3,208 students completing the Flemish medical admission test. Upon completion of the test, applicants were asked to complete measures of test-taking reactions. We received usable answers for 1,750 candidates (66.8% females, 33.2% males; mean age = 18.7 years), yielding a response rate of 54.5%. Nine months after candidates had been admitted, their GPAs were retrieved from the archives of medical universities. This sample demonstrated sufficient power to detect small-strength (f^2 = .06) mediation and moderation effects according to Cohen's (1988) standards (power for mediation analyses = .85; power for moderation analyses = .80).

Measures

Admission test

The medical admission test consisted of a cognitive part (knowledge test) and a noncognitive part (written SJT consisting of physician–patient situations). The total admission test score was a weighted sum of students' test scores on each of these two parts. As only a cognitively oriented criterion measure was available (i.e., GPA), we focused on the job knowledge test and not on the SJT.

Knowledge test

The knowledge test consisted of 40 questions about four sciences (chemistry, physics, mathematics, and biology). Each question had four alternatives. Candidates had 180 minutes to solve the 40 questions. The internal consistency reliability of the knowledge test was .79. Students' knowledge test score was computed by summing their correct answers.

Candidate reactions

Upon completion of the knowledge test, candidates' test anxiety and motivation were each measured with three items from the Test Attitude Survey (Arvey et al., 1990). Example items were "During the testing, I often thought about how poorly I was doing" (test anxiety) and "I wanted to do well on the knowledge test" (test motivation). The internal consistency reliabilities for these scales were .75 and .86, respectively. Candidates' self-efficacy in completing the knowledge test was assessed with three items from Bauer et al. (1998) and had an internal consistency reliability of .82. An example item was "I am confident in my ability to do well on the knowledge test."

In addition to measuring these reactions to the knowledge test, we measured candidates' test beliefs regarding the whole admission exam with four items from the test belief scale of Test Attitude Survey. This measure was administered upon completion of the exam. An example item was "The medical admission test is a good way of selecting people into medical education." The internal consistency of the belief in tests scale was .85. Candidate reaction items were rated on a 5-point Likert-type scale (1 = *strongly disagree* to 5 = *strongly agree*) across all four of our studies.

Performance

Students' GPAs at the end of the first year were retrieved from archival records of all medical universities in Belgium. Given differences between medical universities (e.g., in terms of course content, professors, grading), we standardized the GPAs within each university.

Analytic strategy

Our first two hypotheses were concerned with whether candidate reactions would demonstrate direct (H1) or indirect (H2) effects on job performance. We tested them using the Preacher and Hayes (2004) meditational bootstrapping procedure with 95% bias-corrected confidence estimates and 1,000 bootstrap samples. This procedure enabled the simultaneous examination of direct, indirect, and mediating effects. Our third hypothesis was concerned with whether candidate reactions would moderate the relation between test scores and job performance. We tested this with moderated multiple regression (MMR) models wherein all predictors were standardized. These models involve testing for slope equality between subgroups (e.g., candidates with positive or negative reactions; Bartlett, Bobko, Mosier, & Hannan, 1978; Cleary, 1968; Meade & Tonidandel, 2010). The variables were entered in three steps. Test scores were entered in Step 1, reactions were entered in Step 2, and the interaction between test scores and reactions was entered in Step 3. Slope differences were examined by comparing the model containing both test scores and test reactions to the model containing test scores, test reactions, and their interaction.

Results

Means, standard deviations, and intercorrelations of Study 1 variables are presented in Table 2. Results of the direct (H1) and indirect (H2) tests are presented in Table 3. As indicated in Table 2, test-taking motivation (β =

.14, $p < .05$) and test self-efficacy (β = .11, $p < .05$) exhibited significant direct effects on job performance. However, no significant direct effects were obtained for test-taking anxiety (β = −.05, *ns*) or belief in tests (β = .04, *ns*). These results provide partial support for H1. These findings also provide support for the hypothesis that the relation between candidate reactions and job performance would be stronger when candidate reactions are dispositional (H1a). That is, two of the three dispositional reactions were significantly related to job performance (test-taking motivation and self-efficacy), whereas the situational reaction (belief in tests) was not.

Table 2
Descriptive Statistics and Correlations in Study 1

Variable	*M*	*SD*	1	2	3	4	5	6
Test reactions								
1. Knowledge test anxiety	2.74	0.90						
2. Knowledge test motivation	4.31	0.72	.09**					
3. Knowledge test self-efficacy	3.40	0.80	−.37**	.10**				
4. Admissions test beliefs	2.56	1.00	−.33**	−.07**	.18**			
Test performance								
5. Knowledge test	8.59	3.64	−.33**	.19**	.32**	.30**		
6. Admissions test	18.29	5.23	−.31**	.20**	.27**	.27**	.94**	
Criterion								
7. GPA	0.00	1.00	−.05	.12**	.11*	.04	.34**	.34**

Note. With the exception of GPA ($N = 530$), the N for all other variables was 1,720. Anxiety, motivation, self-efficacy, and test beliefs were rated on a 5-point scale. The maximum on the knowledge test was 20, whereas the maximum on the whole admissions test was 40. Grade point average (GPA) was standardized within university. *SD* = standard deviation.
* $p < .05$. ** $p < .01$.

Table 3
Direct and Indirect Effects of Test Reactions on Test Scores and GPA for Study 1

	DV: GPA	
Predictor	β	*SE*
IV = anxiety, mediator = knowledge test		
Direct effect (c path)	−.05	.05
IV to mediator (a path)	−.18**	.02
Mediator to DV (b path)	.72**	.08
Mediation effect (c^1 path)	.08	.05
Indirect effect (ab path)	−.13**	.02
IV = motivation, mediator = knowledge test		
Direct effect (c path)	.14**	.05
IV to mediator (a path)	.08**	.03
Mediator to DV (b path)	.65**	.08
Mediation effect (c^1 path)	.09*	.05
Indirect effect (ab path)	.05**	.02
IV = self-efficacy, mediator = knowledge test		
Direct effect (c path)	.11*	.05
IV to mediator (a path)	.25**	.02
Mediator to DV (b path)	.72**	.08
Mediation effect (c^1 path)	−.07	.05
Indirect effect	.18**	.03
IV = test beliefs, mediator = admissions test		
Direct effect (c path)	.04	.04
IV to mediator (a path)	.09**	.02
Mediator to DV (b path)	.78**	.10
Mediation effect (c^1 path)	−.03	.04
Indirect effect (ab path)	.07**	.02

Note. $N = 530$ to 537. GPA = grade point average; *SE* = standard error; DV = dependent variable; IV = independent variable.
* $p < .05$. ** $p < .01$

Turning to the tests of indirect effects (H2), the findings in Table 3 indicate partial support for H2a, which proposed that test performance would mediate relations between dispositional candidate reactions and job performance. Test scores significantly mediated the relation between test-taking motivation and job performance (β for mediated effect = .09, $p < .05$) but did not demonstrate evidence of mediation for test-taking anxiety (β = −.08, *ns*), self-efficacy (β = −.07, *ns*), or test beliefs (β = −.03, *ns*). Instead, a significant indirect effect of the relation between test reactions and job performance (via test scores) was found for test anxiety (β = −.13, $p < .01$), self-efficacy (β = .18, $p < .01$), and test beliefs (β = .07, $p < .01$). This latter finding provides partial support for H2b, which proposed that situational-based reactions (i.e., test beliefs) would indirectly affect performance.

Next, we conducted MMR analyses to test whether candidate reactions moderated the relation between test scores and job performance (H3), as well as to see whether candidate reactions have a stronger effect on the relation between test performance and job performance when they are situational than when they are dispositional (H3a). Results are presented in Table 4. As shown, test-taking reactions had no significant effect on criterion-related validity ($\Delta R^2 = .00$, *ns*). Thus, we found no support for H3 or H3a.

Table 4
Moderating Effects of Candidate Reactions for the Test Performance–GPA Relationship in Study 1

Predictor	DV: GPA			
	b	SE_b	R^2	ΔR^2
IV = knowledge test, moderator = test anxiety				
Step 1: Test scores	.67**	.08	.12	
Step 2: Add anxiety	.08	.05	.12	
Step 3: Add Test × Anxiety	.02	.08	.12	.00
IV = knowledge test, moderator = test motivation				
Step 1: Test scores	.67**	.08	.12	
Step 2: Add motivation	.09	.05	.12	
Step 3: Add Test × Motivation	.06	.10	.12	.00
IV = knowledge test, moderator = self-efficacy				
Step 1: Test scores	.66**	.08	.12	
Step 2: Add self-efficacy	−.07	.05	.12	
Step 3: Add Test × Self-Efficacy	.08	.08	.12	.00
IV = admissions test, moderator = test beliefs				
Step 1: Test scores	.77**	.09	.11	
Step 2: Add test beliefs	−.03	.04	.11	
Step 3: Add Test × Test Beliefs	.10	.10	.11	.00

Note. $N = 530$. All variables were standardized prior to entering them in the regression. The table displays unstandardized regression coefficients (shown as *b*). ΔR^2 = the change in variance accounted for when comparing the model containing test scores and test reactions to the model containing test scores, test reactions, and their interaction. GPA = grade point average; DV = dependent variable; IV = independent variable; *SE* = standard error.
** $p < .01$.

Study 2

Method

Participants and procedure

The sample included 152 job applicants selected for professional sales positions in two large multinational corporations. Forty percent of the sample was female. The primary responsibilities of the sales professionals were to represent products and services to prospective customers and to conduct business-to-business selling activities. These activities included devising sales approaches and solutions, marshaling internal support, analyzing sales opportunities, building relationships, and maintaining technical/professional knowledge about the products they represented.

Each applicant completed a multi-item selection inventory, followed by a questionnaire that assessed their perceptions of justice with respect to the tests and the testing process. The selection inventory and questionnaire were completed on computers at the respective worksites. Approximately one year after applicants had been hired, objective sales performance data were obtained. Although this sample did not demonstrate sufficient power to detect small-strength (f^2 = .06) effects according to Cohen's (1988) standards, it did demonstrate sufficient power to detect medium-strength (f^2 = .15) mediation and moderation effects (power for mediation analyses = .95; power for moderation analyses = .93).

Measures

Selection inventory

The selection inventory was developed based on standard test development procedures, including a critical incidents based job analysis and content validity judgments by subject matter experts. The inventory contained a total of 110 items, which were grouped into three types: biodata, situational judgment, and personality. The biodata portion of the inventory (16 items) asked questions about work experiences (e.g., implementing one's own ideas). The situational judgment portion of the inventory (28 items) presented a series of work-related situations (e.g., interacting with a new prospective customer, resolving a complaint from an angry customer) for which candidates were asked to identify the best course of action from a set of alternatives. The personality portion of the inventory (66 items) assessed nine traits related to the sales roles (e.g., achievement orientation). To facilitate decision making, we combined all items into an overall test score. The internal consistency reliability of the selection inventory was .84.

Candidate reactions

Procedural justice was measured with six items adapted from the Selection Procedural Justice Scale (SPJS; Bauer et al., 2001). These items were drawn from several of the subscales, including job relatedness (e.g., "This inventory measured skills and capabilities relevant to the job in question") and chance to perform (e.g., "The inventory provided an opportunity for me to demonstrate my skills and abilities"). The coefficient alpha for this scale was .84.

Job performance

Performance was measured with an objective measure of employee productivity. The measure reflected the dollar amount of sales relative to an employee's sales target.

Results

Results indicated that procedural justice (M = 4.27, SD = .49) was significantly related to test performance (r = .16, $p < .05$) but not to job performance ($r = -.04$, *ns*). Further, test performance and job performance demonstrated a significant correlation (r = .24, $p < .01$).

We followed the same approach as in Study 1 to test the direct (H1) and indirect (H2) effects of test reactions on job performance. Findings indicated that perceptions of procedural justice did not have a significant direct effect

on job performance (β = −.08, *ns*). Thus, H1 was not supported. However, consistent with H2b, the indirect effect of perceptions of justice on job performance (via test scores) was significant (β =.09, $p < .01$). In contrast, there was no evidence of a mediating effect, which supports H2b.

We conducted MMR analyses to assess the extent to which perceptions of justice moderated relations between test performance and job performance (H3). Results are presented in Table 5. As shown, the interaction between reactions and test scores did not explain additional variance in job performance beyond the main effect of the test scores ($\Delta R^2 = .01$, *ns*). Thus, consistent with Study 1, candidate reactions did not affect the criterion-related validity of test scores in relation to job performance.

Table 5
Moderating Effects of Procedural Justice for the Test Performance–Job Performance Relationship in Study 2

Predictor	DV: Job performance			
	b	SE_b	R^2	ΔR^2
IV = selection inventory, moderator = procedural justice				
Step 1: Test scores	.25**	.08	.06**	
Step 2: Add justice	−.08	.08	.07**	
Step 3: Add Test Score × Justice	−.13	.08	.08**	.01

Note. $N = 147$. The table displays unstandardized regression coefficients (shown as *b*). ΔR^2 = the change in variance accounted for when comparing the model containing test scores and test reactions to the model containing test scores, test reactions, and their interaction. DV = dependent variable; IV = independent variable; *SE* = standard error.
** $p < .01$.

Study 3

Method

Participants and procedure

This study examined a sample of 161 applicants for sales positions at several branches of a large financial institution in South America. Sixty-two percent of the sample was female. The primary responsibilities of the sales position were to conduct interviews with clients, build and maintain strong relationships, effectively sell packages to clients, and stay up-to-date on products and packages offered by the company. Each applicant received an e-mail with a link to a web-based sales assessment that included a personality inventory and a business-related cognitive ability test. Upon completion of the assessment, applicants were asked to complete measures of candidate reactions. After selected employees had been on the jobfor 1 year, supervisor ratings were collected from each employee's immediate supervisor. Although this sample did not demonstrate sufficient power to detect small-strength ($f^2 = .06$) effects according to Cohen's (1988) standards, it did demonstrate sufficient power to detect medium-strength ($f^2 = .15$) mediation and moderation effects (power for mediation analyses = .90; power for moderation analyses = .84).

Measures

The personality inventory and business-related cognitive ability test were developed based on the results of a comprehensive job analysis. The first stage of the job analysis involved focus groups with employees and the

collection of background information on the job. The second stage comprised interviews with employees and supervisors, who were asked to identify the KSAOs required to be successful on the job.

Personality inventory

The personality inventory measured four KSAOs identified as important from the job analyses. The relevant scales were labeled Achievement Orientation (21 items, $\alpha = .88$), Adaptability (16 items, $\alpha = .78$), Interpersonal (48 items, $\alpha = .90$), and Time Management (14 items, $\alpha = .81$).

Business-related cognitive ability test

The business-related cognitive ability test (cf. Hattrup, Schmitt, & Landis, 1992) was developed to measure the ability to analyze and interpret sales data. It contained 15 questions that asked applicants to interpret sales charts and graphs, compute sales figures, and solve reasoning problems. Items were either fill in the blank computations or multiple-choice reasoning problems. Each item was scored right or wrong. The coefficient alpha for the test was .70.

Candidate reactions

Procedural justice was measured with six items adapted from the SPJS (Bauer et al., 2001; $\alpha = .76$). These items were drawn from several of the subscales, including job relatedness (e.g., "Applicants who do well on this assessment will probably do well in sales role") and chance to perform (e.g., "Applicants who do well on this assessment will probably do well in sales role").

Job performance

Supervisors rated employees on 20 items using a 1 (s*trongly disagree*) to 7 (s*trongly agree*) scale. Items reflected four dimensions: Drive (e.g., "This employee clearly demonstrates a strong desire to succeed"; 5 items, $\alpha = .91$); Flexibility (e.g., "This employee adapts to changing demands"; 5 items, $\alpha = .89$), People-Oriented (e.g., "This employee empathizes and understands people"; 5 items, $\alpha = .86$), and Accountability (e.g., "This employee follows through on responsibilities"; 5 items, $\alpha = .88$). We also averaged the dimension ratings to form an overall performance composite ($\alpha = .97$). A confirmatory factor analysis indicated that the four-factor structure provided a good fit to the data ($\chi(48)^2 = 103.4$, $p < .01$; comparative fit index [CFI] = .97; root-mean-square error of approximation [RMSEA] = .08) and fit significantly better than a one-factor model ($\Delta\chi(6)^2 = 18.6$, $p < .01$).

Results

Means, standard deviations, and intercorrelations are presented in Table 6. For this study, we were able to test our conceptual matching hypothesis (H1b) by matching the four personality test scales (Achievement, Adaptability, Interpersonal, and Time Management) with the four dimensions of job performance (Drive, Flexibility, People-Oriented, and Accountability). Findings are presented in Table 7. As illustrated, procedural justice did not demonstrate a significant direct effect on any of the job performance measures (βs = −.06 to −.01, *ns*). Thus, H1 and H1b were not supported. There was also no evidence of a mediating or an indirect effect (of the relation between perceptions of justice and job performance) for personality test scores or cognitive ability test scores (βs = −.09 to .02, *ns*). Thus, H2b was not supported.

Table 6
Descriptive Statistics and Correlations in Study 3

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
Test reaction												
1. Procedural justice	3.99	0.57										
Test performance												
2. Personality scale: Achievement	0.00	1.00	.26**									
3. Personality scale: Adaptability	0.00	1.00	.17*	.51**								
4. Personality scale: Interpersonal	0.00	1.00	.16	.38**	.71**							
5. Personality scale: Time Management	0.00	1.00	.26**	.43**	.66**	.66**						
6. Business-related cognitive ability test	0.00	1.00	−.05	.06	.09	.25*	.14					
Job performance												
7. Job performance: Drive	0.00	1.00	−.01	.12	.15*	.18*	.15*	.11				
8. Job performance: Flexibility	0.00	1.00	−.07	.10	.15*	.12	.08	.13	.75**			
9. Job performance: People-Oriented	0.00	1.00	−.04	.04	.07	.08	.01	.16*	.66**	.74**		
10. Job performance: Accountability	0.00	1.00	−.02	.03	.14	.08	.10	.16*	.79**	.83**	.73**	
11. Job performance: Overall	0.00	0.00	−.03	.08	.13	.11	.08	.15	.88**	.92**	.87**	.93**

Note. $N = 108$ to 161. *SD* = standard deviation. Procedural justice was rated on a 5-point scale. The data were obtained from multiple branches; thus, standardized scores were used.
* $p < .05$. ** $p < .01$.

Table 7
Direct and Indirect Effects of Procedural Justice on Test Scores and Job Performance for Study 3

	DV: Job performance[a]	
Predictor	β	*SE*
IV = procedural justice, mediator = personality: Achievement		
Direct effect (c path)	−.01	.10
IV to mediator (a path)	.27**	.10
Mediator to DV (b path)	.02	.10
Mediation effect (c^I path)	−.01	.10
Indirect effect (ab path)	.01	.03
IV = procedural justice, mediator = personality: Adaptability		
Direct effect (c path)	−.06	.09
IV to mediator (a path)	.16	.09
Mediator to DV (b path)	.15	.09
Mediation effect (c^I path)	−.09	.09
Indirect effect (ab path)	.02	.02
IV = procedural justice, mediator = personality: Interpersonal		
Direct effect (c path)	−.04	.09
IV to mediator (a path)	.14	.09
Mediator to DV (b path)	.08	.10
Mediation effect (c^I path)	−.05	.09
Indirect effect (ab path)	.01	.02
IV = Procedural justice, mediator = personality: Time Management		
Direct effect (c path)	−.02	.09
IV to mediator (a path)	.24**	.09
Mediator to DV (b path)	.10	.10
Mediation effect (c^I path)	−.04	.10
Indirect effect (ab path)	.02	.03

Note. $N = 107$ to 108. *SE* = standard error; DV = dependent variable; IV = independent variable.
[a] Personality scales were conceptually matched to the relevant job performance scales; Achievement was examined with the job performance Drive scale, Adaptability was examined with the job performance Flexibility scale, Interpersonal was examined with the job performance People-Oriented scale, and Time Management was examined with the job performance Accountability scale.

Next, we examined whether perceptions of justice moderated relations between job performance and each of the four personality test scales and the ability test. Consistent with Studies 1 and 2, candidate reactions did not affect the criterion-related validity of any of the personality scales or the cognitive ability test (see Table 8). Further, the magnitude of effects was trivial, with ΔR^2 values ranging from .00 to .02. These findings provide no support for H3 or H3a. Finally, the absence of different effects for the four dimensions of job performance suggests little support for H3b (i.e., that moderation would be stronger when test reactions and job performance are conceptually matched).

Table 8
Moderating Effects of Test Reactions on the Test Performance–Job Performance Relations in Study 3

Predictor	DV: Job performance[a]			
	b	SE_b	R^2	ΔR^2
IV = personality Achievement, DV = job performance Drive				
Step 1: Achievement scale	.02	.10	.00	
Step 2: Add justice	−.02	.10	.00	
Step 3: Add Achievement × Justice	−.09	.11	.01	.01
IV = personality Adaptability, DV = job performance Flexibility				
Step 1: Adaptability scale	.13	.09	.02	
Step 2: Add justice	−.09	.09	.03	
Step 3: Add Adaptability × Justice	.04	.10	.03	.00
IV = personality Interpersonal, DV = job performance People-Oriented				
Step 1: Interpersonal scale	.07	.09	.01	
Step 2: Add justice	−.05	.09	.01	
Step 3: Add Interpersonal × Justice	−.06	.11	.01	.00
IV = personality Time Management, DV = job performance Accountability				
Step 1: Time Management scale	.09	.10	.00	
Step 2: Add justice	−.04	.10	.01	
Step 3: Add Time Management × Justice	.09	.10	.01	.00
IV = cognitive ability, DV = job performance Overall				
Step 1: Cognitive ability test	.11	.10	.01	
Step 2: Add justice	−.01	.09	.01	
Step 3: Add Test Score × Justice	.16	.11	.03	.02

Note. $N = 108$. SE = standard error; DV = dependent variable; IV = independent variable. The table displays unstandardized regression coefficients (shown as *b*). ΔR^2 = change in variance accounted for when comparing the model containing test scores and test reactions to the model containing test scores, test reactions, and their interaction.

[a] Personality scales were conceptually matched to the relevant job performance scales; Achievement was examined with the job performance Drive scale, Adaptability was examined with the job performance Flexibility scale, Interpersonal was examined with the job performance People-Oriented scale, and Time Management was examined with the job performance Accountability scale.

Study 4

Method

Participants and procedure

The sample comprised 2,959 employees working for a large North American organization. Seventy percent of the sample was female. Participants held full-time jobs either as customer service representatives (N = 1,409) or as product technicians (N = 1,550). The primary responsibilities of the customer service representatives are to describe, recommend, and help customers with product lines. Service representatives also process payments and deal with all customer service needs. The primary responsibilities of the product technicians are to develop and produce products, maintain proper storage and security of products, and manage product inventory.

Each participant completed three selection tests: a work sample test, a SJT, and a personality test. Candidates also completed a battery of measures that assessed their reactions after completing the set of selection procedures. The tests were completed on computers at the respective worksites. Concurrently, job performance ratings were obtained by having the immediate supervisors of each participant complete an on-line performance evaluation that was developed for the current study and used for research purposes only. This sample demonstrated sufficient power to detect small-strength (f^2 = .06) mediation and moderation effects according to Cohen's (1988) standards (power for mediation and moderation analyses = .99).

Measures

The three assessments were developed based on the results of comprehensive job analyses. The first stage of the job analysis involved focus groups with employees and the collection of background information on the job. This resulted in detailed lists of the tasks and KSAOs that were required for each job. The second stage comprised a large-scale job analysis survey conducted with employees to measure the jobs more precisely (e.g., importance and frequency for tasks, and importance and needed at entry for KSAOs). Three separate forms of content validity evidence were also evaluated for each selection procedure: (a) a rational analysis of content based on testing guidelines, (b) linkages between the job analysis task and KSAOs, and (c) subject matter expert (SME) ratings of the job relatedness of the questions within each assessment.

Work sample test

The work sample test consisted of 20 digital pictures that represented the tasks on the job. A sample picture for the customer service representatives would be of a product label that the employee must read and explain to the customer. A sample picture for the product technician would be two products and their ingredients that must be compared and contrasted to one another. A multiple-choice question was written for each picture. Internal consistency reliability estimates were .72 for the customer service representative simulation and .64 for the product technician simulation.

Situational judgment test

The SJT presented a series of work-related situations for which candidates were asked to identify the best course of action from a set of alternatives. During the job analysis, a large sample of critical incidents was collected through surveys of store employees. This information was used to create a series of multiple-choice items that covered a wide range of actual situations that occur on the job. A panel of SMEs rated each item to facilitate item selection and to provide evidence of content validity. The SMEs were also asked to indicate the best answer and to make edits and corrections to the questions and answers.

A total of 20 items were selected for the final SJT, which had an internal consistency reliability of .48. Modest reliability estimates such as this are common for SJTs, as they tend to measure multiple constructs and thus may not be as homogeneous as other tests (McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001; Schmitt & Chan, 2006).

Personality test

The personality test was designed to provide a content valid measure of personality-related attributes relevant to the job. First, the KSAOs identified from the job analysis were translated into traits commonly used in personality research. Next, items were written for each trait. An initial set of items was pilot-tested on a sample of 78 undergraduate students at a large university to evaluate the psychometric properties of the items and to make necessary adjustments. A panel of SMEs then evaluated the content validity of the revised items. The final version of the personality test contained three subscales, comprising eight items each: Goal Orientation (α = .73; e.g., "I use goal setting more often than other people"), Composure (α = .74; "It is hard to be patient at the end of a long workday"), and Helping (α = .72; "I like to be available to others in case they need me"). Items were also combined to create an overall score (α = .81).

Candidate reactions

Five scales assessed participant reactions to the tests. Two items from the Test Attitude Survey (Arvey et al., 1990) were adapted to assess test-taking anxiety (α = .78 for each of the three tests; e.g., "This test made me nervous, which had a negative impact on my answers"). Test-taking motivation was measured with three items adapted from the Motivation subscale of the Test Attitude Survey (α = .74 for each of the three tests; e.g., "I tried my best on this test"). Procedural justice was measured with four content/predictive validity items from the SPJS (Bauer et al., 2001; α = .81 to .83 across the three tests; e.g., "The content of this test is related to the job").

Job performance

Supervisors evaluated the performance of each candidate on 10 dimensions identified by the job analysis. The dimensions reflected a combination of Task performance (5 items; e.g., "Rate this employee on their level of job knowledge"; α = .91) and Interpersonal performance (5 items; e.g., "Rate this employee on their interpersonal skills"; α = .91). Supervisors rated each dimension on a Likert-type scale (1 = *unacceptable*; 5 = *outstanding*). A confirmatory factor analysis indicated that the two-factor structure provided a good fit to the data ($\chi(26)^2 = 522.2$, $p < .01$; CFI = .98; RMSEA = .08) and fit significantly better than a one-factor model ($\Delta\chi(27)^2 = 56.88$, $p < .01$). In the majority of cases (88%), the performance ratings were made by two managers, and the intraclass correlation coefficient (ICC, 2) with respect to their rating was .65. Thus, we averaged the two sets of scores across the 10 dimensions.

Results

The means, standard deviations, and intercorrelations of the study variables for the customer service representative and product technician jobs are presented in Tables 9 and 10, respectively. For this study, we were able to test our conceptual matching hypothesis (H1b) by matching the test reactions to the relevant personality and job performance scales. In particular, test-taking anxiety was matched to the personality Composure scale and the Task job performance scale; motivation was matched to the personality Goal Orientation scale and the Task job performance scale; and justice was matched to the personality Helping scale and the Interpersonal job performance scale.

Table 9
Descriptive Statistics and Correlations for Customer Service Representatives in Study 4

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Anxiety																	
1. Work sample	2.12	0.94	—														
2. SJT	1.99	0.91	.88	—													
3. Personality	1.99	0.93	.87	.95	—												
Motivation																	
4. Work sample	4.15	0.63	−.40	−.39	−.38	—											
5. SJT	4.12	0.64	−.37	−.40	−.38	.89	—										
6. Personality	4.13	0.63	−.36	−.39	−.38	.87	.92	—									
Justice																	
7. Work sample	4.25	0.68	−.27	−.30	−.28	.43	.41	.40	—								
8. SJT	4.16	0.68	−.24	−.27	−.26	.38	.40	.39	.79	—							
9. Personality	4.12	0.72	−.19	−.23	−.22	.34	.35	.38	.69	.83							
Test performance										.15							
10. Work sample	82.11	11.50	−.25	−.24	−.24	.25	.24	.22	.21		.32						
11. SJT	72.54	12.86	−.15	−.18	−.17	.15	.13	.15	.10	.08	.06	.42					
12. Personality: Composure	4.21	0.49	−.22	−.24	−.24	.28	.26	.28	.25	.26	.27	.09	.15				
13. Personality: Goal Orientation	4.00	0.50	−.18	−.21	−.21	.25	.25	.25	.26	.29	.26	.10	.19	.57			
14. Personality: Helping	4.26	0.45	−.13	−.17	−.16	.25	.26	.25	.26	.28	.27	.08	.17	.58	.58		
Job performance																	
15. Task-based	3.15	0.68	−.06	−.06	−.06	.00	−.01	−.01	.02	.01	−.01	.14	.21	.00	.06	.01	
16. Interpersonal-based	2.96	0.61	−.03	−.03	−.04	−.02	−.01	.00	−.02	−.01	.00	.09	.19	.03	.07	.06	.82

Note. N = 1,216 to 1,350. Correlations ≥ .09 are significant at $p < .05$; correlations ≥ .10 are significant at $p < .01$. The values for the work sample, SJT, and personality test within each variable category reflect candidates' perceptions of anxiety, motivation, and justice. Anxiety, motivation, justice, personality and job performance were rated on a 5-point scale. The minimum value on the work sample and SJT tests was 0, whereas the maximum was 100. *SD* = standard deviation; SJT = situational judgment test.

Table 10
Descriptive Statistics and Correlations for Product Technicians in Study 4

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Anxiety																	
1. Work sample	1.94	0.88	—														
2. SJT	1.85	0.85	.88	—													
3. Personality	1.86	0.85	.86	.95	—												
Motivation																	
4. Work sample	4.18	0.66	−.44	−.45	−.44	—											
5. SJT	4.13	0.66	−.43	−.43	−.43	.91	—										
6. Personality	4.13	0.66	−.42	−.42	−.44	.89	.95	—									
Justice																	
7. Work sample	4.15	0.72	−.26	−.29	−.27	.45	.41	.39	—								
8. SJT	3.97	0.73	−.24	−.27	−.25	.36	.39	.37	.76	—							
9. Personality	3.95	0.76	−.18	−.20	−.20	.32	.37	.39	.67	.80							
Test performance																	
10. Work sample	86.48	11.50	−.24	−.22	−.23	.16	.14	.13	.07	.00	−.02	—					
11. SJT	75.96	11.84	−.18	−.18	−.17	.12	.14	.14	.12	.09	.05	.30	—				
12. Personality: Composure	4.00	0.50	−.19	−.21	−.22	.27	.28	.29	.23	.27	.28	.07	.17	—			
13. Personality: Goal Setting	4.22	0.48	−.18	−.20	−.21	.27	.29	.30	.22	.27	.28	.08	.17	.62	—		
14. Personality: Helping	4.29	0.44	−.15	−.17	−.18	.26	.27	.27	.24	.28	.27	.09	.13	.59	.58	—	
Job performance																	
15. Task based	3.02	0.62	−.11	−.12	−.12	.02	.03	.03	−.01	−.03	−.03	.17	.17	.04	.04	.04	
16. Interpersonal based	3.17	0.68	−.09	−.11	−.12	.02	.02	.02	.04	.05	.04	.11	.17	.11	.06	.08	.80

Note. N = 1,390 to 1,498. Correlations ≥ .05 are significant at $p < .05$; correlations ≥ .07 are significant at $p < .01$. The values for the work sample, SJT, and personality test within each variable category reflect candidates' perceptions of anxiety, motivation, and justice. Anxiety, motivation, justice, personality and job performance were rated on a 5-point scale. The minimum value on the work sample and SJT tests was 0, whereas the maximum was 100. *SD* = standard deviation; SJT = situational judgment test.

Tables 11 and 12 present the tests of the direct (H1) and indirect (H2) effects when the test reactions were conceptually matched to the relevant personality and job performance scales. As indicated, test-taking anxiety exhibited a direct effect on job performance for the product technician jobs across all three tests (βs ranged from −.21 to −.24, all *p*s < .01). In contrast, test-taking anxiety did not demonstrate a significant direct effect on performance for the customer service representative jobs (βs ranged from −.08 to −.09, all *ns*). Similarly, test-taking motivation and justice did not demonstrate significant direct effects on performance for either job or any of the three tests (βs ranged from −.04 to.07, all *ns*). Thus, we found partial support for H1 (direct effect of reactions) and H1a (stronger direct relations for dispositional reactions). However, all tests of H1 indicated a similar pattern of findings when test reactions were conceptually matched to relevant scales versus when they were not conceptually matched to their relevant scales. Thus, no support for H1b, which predicted stronger results for conceptually matched scales, was obtained.

Table 11
Direct and Indirect Effects of Test Reactions on Test Scores and Job Performance for Customer Service Representatives in Study 4

	Work sample test		SJT		Personality test[a]	
Predictor	β	*SE*	β	*SE*	β	*SE*
Anxiety [a]						
Direct effect (c path)	−.08	.05	−.09	.05	−.08	.05
IV to mediator (a path)	−.13**	.02	−.17**	.03	−.27**	.03
Mediator to DV (b path)	.38**	.07	.39**	.05	.16**	.05
Mediation effect (c^1 path)	−.04	.04	−.02	.05	−.04	.05
Indirect effect (ab path)	−.05**	.01	−.06**	.01	−.04**	.01
Motivation [a]						
Direct effect (c path)	−.03	.05	−.01	.05	−.01	.05
IV to mediator (a path)	.10**	.02	.14**	.03	.36**	.03
Mediator to DV (b path)	.39**	.07	.40**	.05	.21**	.05
Mediation effect (c^1 path)	−.07	.05	−.07	.05	−.08	.05
Indirect effect (ab path)	.04**	.01	.05**	.01	.07**	.02
Justice[a]						
Direct effect (c path)	.00	.05	−.03	.05	−.04	.05
IV to mediator (a path)	.07**	.02	.10**	.03	.33**	.03
Mediator to DV (b path)	.38**	.07	.41	.05	.21**	.05
Mediation effect (c^1 path)	−.03	.05	−.07	.05	−.11*	.05
Indirect effect (ab path)	.04**	.01	.04**	.01	.07**	.02

Note. N = 1,216 to 1,277. SJT = situational judgment test; *SE* = standard error; IV = test reaction; mediator = test score; DV = job performance.

[a] Test reactions were conceptually matched to the relevant personality and job performance scales; test anxiety was examined with the personality Composure scale and the Task job performance scale, test motivation was examined with the personality Goal Setting scale and the Task job performance scale, and justice was examined with the personality Helping scale and the Interpersonal job performance scale.

* $p < .05$. ** $p < .01$.

Table 12
Direct and Indirect Effects of Test Reactions on Test Scores and Job Performance for Product Technicians in Study 4

Predictor	Work sample test		SJT		Personality test[a]	
	β	*SE*	β	*SE*	β	*SE*
Anxiety						
Direct effect (c path)	−.21**	.05	−.22**	.05	−.24**	.05
IV to mediator (a path)	−.13**	.02	−.17**	.03	−.27**	.03
Mediator to DV (b path)	.34**	.07	.34**	.05	.15**	.05
Mediation effect (c^1 path)	−.16**	.05	−.16**	.05	−.20**	.05
Indirect effect (ab path)	−.05**	.01	−.06**	.01	−.04**	.01
Motivation						
Direct effect (c path)	.05	.05	.07	.05	.06	.04
IV to mediator (a path)	.05*	.02	.12**	.02	.31**	.02
Mediator to DV (b path)	.38**	.07	.36**	.05	.20**	.05
Mediation effect(c^1 path)	.03	.05	.03	.05	.00	.05
Indirect effect (ab path)	.02**	.01	.05**	.01	.06**	.02
Justice						
Direct effect (c path)	.02	.05	.03	.05	.02	.04
IV to mediator (a path)	.01	.02	.08**	.08	.30**	.02
Mediator to DV (b path)	.38**	.07	.35**	.05	.22**	.05
Mediation effect (c^1 path)	.01	.05	.00	.04	−.05	.05
Indirect effect (ab path)	.00	.01	.03**	.01	.06**	.02

Note. N = 1,388 to 1,417. SJT = situational judgment test; *SE* = standard error; IV = test reaction; mediator = test score; DV = job performance.
[a] Test reactions were conceptually matched to the relevant personality and job performance scales; test anxiety was examined with the personality Composure scale and the Task job performance scale, test motivation was examined with the personality Goal Setting scale and the Task job performance scale, and justice was examined with the personality Helping scale and the Interpersonal job performance scale.
** $p < .01$.

Concerning the indirect relationships (H2), test scores significantly mediated the relation between test-taking anxiety and job performance for the product technician jobs across all three tests (βs ranged from −.16 to −.20, all *p*s < .01). These results provide partial support for H2a (mediation for dispositional reactions). In all other cases, candidate reactions demonstrated a significant indirect effect on job performance via their influence on test performance (βs ranged from −.06 to .07, all *p*s < .01). The one exception was for justice perceptions of the work sample test among product technicians (β = .00, *ns*). These findings provide support for H2b in that situational test reactions tended to exhibit an indirect effect on job performance.

Results of the MMR regression analyses are presented in Tables 13 and 14. As illustrated, candidate reactions had no significant effect on criterion-related validity for the customer service representative job or the product technician job (ΔR^2 = .00). These findings provide no support for H3 or H3a. Further, the absence of different effects for the two dimensions of job performance suggests no support for H3b (i.e., that moderation would be stronger when test reactions and job performance are conceptually matched). Overall, these findings generally are consistent with Studies 1–3 and provide the strongest support for the hypothesis that candidate reactions exhibit indirect relations with job performance.

Table 13
Moderating Effects of Test Reactions on the Test Performance–Job Performance Relations for the Customer Service Representative Job in Study 4

Predictor	Work sample test			SJT			Personality test		
	b	R^2	ΔR^2	*b*	R^2	ΔR^2	*b*	R^2	ΔR^2
Anxiety									
Step 1: Test scores	.16**	.04**		.13**	.05**		.03	.00	
Step 2: Add anxiety	−.01	.04**		−.01	.05**		−.03	.00	
Step 3: Add Test × Anxiety	−.02	.04**	.00	.01	.05**	.00	.03*	.00	.00
Motivation									
Step 1: Test scores	.16**	.03**		.13**	.05**		.03**	.00	
Step 2: Add motivation	−.02	.03**		−.02	.05**		−.02	.00	
Step 3: Add Test × Motivation	.03	.03**	.00	.00	.05**	.00	−.02	.00	.00
Justice									
Step 1: Test scores	.11**	.01**		.13**	.04**		.06**	.01**	
Step 2: Add justice	−.02	.01**		−.02	.04**		−.02	.01**	
Step 3: Add Test × Justice	.03	.01**	.00	−.01	.04**	.00	−.04*	.01**	.00

Note. N = 1,216 to 1,281. The table displays unstandardized regression coefficients (*b*). Test reactions were conceptually matched to the relevant personality and job performance scales; test anxiety was examined with the personality Composure scale and the Task job performance scale, test motivation was examined with the personality Goal Setting scale and the Task job performance scale, and justice was examined with the personality Helping scale and the Interpersonal job performance scale.ΔR^2 = change in variance accounted for when comparing the model containing test scores and test reactions to the model containing test scores, test reactions, and their interaction. SJT = situational judgment test.
* $p < .05$. ** $p < .01$.

Table 14
Moderating Effects of Test Reactions on the Test Performance–Job Performance Relations for the Product Technician Jobs in Study 4

Predictor	Work sample test			SJT			Personality test		
	b	R^2	ΔR^2	*b*	R^2	ΔR^2	*b*	R^2	ΔR^2
Anxiety									
Step 1: Test scores	.16**	.03**		.11**	.03**		.04**	.01**	
Step 2: Add anxiety	−.05**	.04**		−.06**	.04**		−.08**	.02**	
Step 3: Add Test × Anxiety	.04*	.04**	.00	.02	.04**	.00	.03	.02**	.00
Motivation									
Step 1: Test scores	.16**	.03**		.11**	.03**		.04*	.00	
Step 2: Add motivation	.01	.03**		.01	.03**		.01	.00	
Step 3: Add Test × Motivation	.00	.03**	.00	.01	.03**	.00	−.02	.00	.00
Justice									
Step 1: Test scores	.09**	.01**		.12**	.03**		.06**	.01**	
Step 2: Add justice	.03	.01**		.02	.03**		.01	.01**	
Step 3: Add Test × Justice	.04	.01**	.00	.02	.03**	.00	−.01	.01**	.00

Note. N = 1,395 to 1,424. The table displays unstandardized regression coefficients (*b*). Test reactions were conceptually matched to the relevant personality and job performance scales; test anxiety was examined with the personality Composure scale and the Task job performance scale, test motivation was examined with the personality Goal Setting scale and the Task job performance scale, and justice was examined with the personality Helping scale and the Interpersonal job performance scale. ΔR^2 = change in variance accounted for when comparing the model containing test scores and test reactions to the model containing test scores, test reactions, and their interaction. SJT = situational judgment test.
* $p < .05$. ** $p < .01$.

Summary of Overall Results

Table 15 provides a summary of results across the four studies. Overall, we found considerable support for the direct effect hypotheses for dispositional reactions (H1a). There was also partial support for a mediating effect of test scores on the relation between test reactions and job performance for dispositional reactions (H2a) and an indirect effect of test reactions on job performance (via test scores) for situational reactions(H2b). Finally, we found no support for our conceptual matching hypotheses (H1b, H3c) or for the moderating effect of candidate reactions on the relations between test performance and job performance (H3, H3a).

Table 15
Summary of Findings Across the Four Studies

Effect type	Study 1	Study 2	Study 3	Study 4
1: Direct effect				
H1: Reactions have a direct effect on job performance	Partial support –motivation –self-efficacy	Not supported	Not supported	Partial support –anxiety (PTs only)
H1a: Effect is stronger when reactions are dispositional	Partial support –motivation –self-efficacy	Not applicable	Not applicable	Partial support –anxiety (PTs only)
H1b: Effect is stronger when conceptually matched	Not applicable	Not applicable	Not supported	Not supported
2: Indirect effect				
H2a: When reactions are dispositional, test scores mediate	Partial support –motivation	Not applicable	Not applicable	Partial support –anxiety (PTs only)
H2b: When reactions are situational, they exhibit an indirect effect	Partial support –test beliefs –anxiety[a] –self-efficacy[a]	Supported –justice	Not supported	Partial support –justice –anxiety[a] (CSRs only) –motivation[a]
3: Moderating effect				
H3: Test reactions moderate test performance–job performance	Not supported	Not supported	Not supported	Not supported
H3a: Effect is stronger when reactions are situational	Not supported	Not applicable	Not supported	Not supported
H3b: Effect is stronger when conceptually matched	Not applicable	Not applicable	Not supported	Not supported

Note. H = hypothesis; PTs = product technicians; CSRs = customer service representatives.
[a]A significant effect was obtained but not predicted.

General Discussion

This research contributes to the literature on candidate reactions to selection procedures by examining the effects of reactions on an understudied, yet vitally important, criterion variable: job performance. Drawing from theory and past research, we provide a conceptual framework (see Table 1) that delineates three ways in which candidate reactions may influence job performance and affect how selection procedures predict job performance. We tested the hypotheses from this framework using data from a diverse set of contexts, samples, and measures. The pattern of results yielded three key findings. First, when candidate reactions are dispositional in nature, test scores mediate the relation between test reactions and job performance. Second, when candidate reactions are situational in nature, they have an indirect effect on job performance (through their relation with test performance). Third, irrespective of their nature, candidate reactions have little or no effect on the relation between test performance and performance on the job. In other words, reactions did not appear to have any influence (positive or negative) on the criterion-related validity of the various selection procedures we examined. These findings were consistent across four different jobs, five types of candidate reactions, six selection procedures, two study contexts, three continents, and two types of respondents.

Implications for Theory, Research, and Practice

This study contributes to theory by advancing a conceptual framework for considering potential relations among candidate reactions, test performance, and job performance, and it helps to answer calls to integrate the variety of applicant reactions (Ryan & Ployhart, 2000). Results support the notion that the theoretical link between *dispositional* test reactions and performance is more proximal than the link between *situational* test reactions and job performance. This provides assimilation and reinforcement of three relevant theoretical perspectives: affective events theory (Weiss & Cropanzano, 1996), attitude–behavior relations (Ajzen & Fishbein, 1977), and models of job performance (e.g., Campbell et al., 1993). These theories share the assumption that attitude–behavior relations are strongest when they are proximal to each other in terms of duration and focus. Our findings support this proposition and have notable implications for theoretical models of candidate reactions, which often include job

performance as a key outcome variable (e.g., Gilliland, 1993, 1994; Hausknecht et al., 2004; Ryan & Ployhart, 2000). Moving forward, it would be valuable for future studies to conduct longitudinal research that distinguishes between dispositional versus situational reactions. Future work that considers other conceptual and empirical distinctions among applicant perceptions also would be valuable.

In contrast to our dispositional hypotheses, the conceptual matching hypotheses were not supported. This may be a function of the specific reactions and outcomes we were able to assess. There are a multitude of potential matches that could occur between candidate reactions, test scores, and job performance, and it would be worthwhile for future research to explore a broader range of possibilities. For example, it would be interesting to examine relations among perceptions of procedural justice, scores on integrity tests, and measures of counterproductive performance. Ultimately, studies that involve more detailed levels of conceptual matching are likely to prove advantageous.

From a practical perspective, a key takeaway is that candidate reactions do not appear to affect the criterion-related validity of selection procedures for the criterion of job performance. One implication of this finding is that reactions to selection procedures that are prone to more negative reactions, such as personality tests (Anderson & Witvliet, 2008; Hausknecht et al., 2004), might not adversely affect criterion-relatedvalidity. At the same time, some of our findings suggest that candidates with negative reactions demonstrate somewhat lower levels of job performance. If so, it may not be in the best interest of organizations to try to reduce applicants' anxiety or to increase their motivation and self-efficacy. Indeed, it may be useful to allow individual differences in these constructs to emerge as true reflections of performance potential, particularly in highly stressful and demanding occupations such as first responders, airline pilots, and taxi drivers. Incorporating stress testing (e.g., stress interviews, stress tolerance scales) into the selection process is also likely to prove advantageous for these occupational groups.

This does not imply, however, that organizations should be insensitive to candidate test reactions. On the contrary, as our results suggest, certain types of reactions can have a direct influence on job performance. Also, reactions have been found to impact other important outcomes, such as organizational attractiveness, recommendation intentions, and job acceptance intentions (Hausknecht et al., 2004). Further, from the candidates' point of view, reactions can negatively affect test performance, which may reduce the possibility of obtaining a desired position. Thus, techniques to minimize test anxiety and to increase perceptions of motivation and justice are important. Training programs for job applicants, such as relaxation techniques (Poppen, 1988) and self-efficacy training (Eden & Aviram, 1993), may prove particularly useful. Test preparation and retesting are also likely to help (Schleicher, Van Iddekinge, Morgeson, & Campion, 2010), as would giving applicants advance information about the selection process and tests (Truxillo et al., 2002).

Limitations and Future Research

Although the present research is characterized by several notable strengths, it also includes certain limitations. To begin, there may be alternative explanations as to why we did not find any evidence that applicant reactions moderate relations between test scores and job performance. However, several aspects of our research would seem to allow for fairly strong inferences about the effects of reactions on criterion-related validity. First, we engaged in triangulation and multiple operationalizations by approaching our research question from several angles (Cortina & Folger, 1998). In doing so, we were able to examine the effects of candidate reactions across different industries, contexts, countries, jobs, types of test reactions, and types of selection procedures. Second, the nonsignificant moderation effects we observed were accompanied by a set of anticipated positive relations. In particular, test reactions were significantly related to test performance, and test performance was significantly related to job performance across all four studies. Third, two of our studies contained large sample sizes (Study 1, $N = 530$; Study 4, $N = 2{,}959$), which enabled us to examine differences in criterion-related validity with enough power to detect extremely small effects (e.g., Aguinis, Culpepper, & Pierce, 2010; Aguinis & Stone-Romero, 1997). Fourth, we examined candidate reactions among both job applicants and current employees. In no case did we find evidence that reactions influenced criterion-related validity, even in the incumbent samples where there

was greater variance in test scores and reactions. Fifth, we found no evidence that range restriction in applicant reactions or test scores contributed to the lack of effects for some of our hypotheses. This is important, and we encourage future research on applicant reactions to consider the possible influence of range restriction.

We also note that applicant reactions were measured after the completion of the tests. Although this approach allows applicants to experience each test before reaction measures are obtained, it also may allow applicants' perceptions of their test performance to affect their reactions to the testing process (e.g., a self-serving bias; Chan, Schmitt, Jennings, et al., 1998). We further acknowledge that it was not possible to provide exhaustive coverage of all possible selection techniques used by organizations. Thus, although we suspect that our findings would receive support using other selection procedures, future research that examines the potential influence of candidate reactions on the validity of other selection procedures, such as employment interviews, may prove useful. It would also be worthwhile for future research to consider whether different groups of individuals experience varying levels and types of test reactions. For example, individuals who have higher levels of education and increased exposure to tests may demonstrate higher test-taking motivation and self-efficacy, as well as lower levels of test anxiety. As a result, candidate reactions may prove less detrimental in high-level white-collar positions and more problematic in blue-collar jobs. Future research should also examine whether test reactions affect criterion-related validity in promotional contexts, as the consequences associated with promotional candidates may be more severe than those associated with job applicants (Ford, Truxillo, & Bauer, 2009). For example, existing employees who feel that promotion tests or processes are unfair may develop negative attitudes, reduce their performance, or even engage in counterproductive workplace behaviors (McCarthy et al., 2009).

Future research should also consider a third boundary condition that may influence relations between candidate reactions and job performance: the extent to which the criterion is proximal (e.g., performance during training) versus distal (e.g., job performance) to the reactions. Considerable theoretical and empirical evidence suggests that relations between two variables are stronger when they are proximal in time (e.g., Harrison, Newman, & Roth, 2006; Krausz, Kozlowsky, & Eiser, 1998). Thus, test reactions may be more likely to affect proximal measures of job performance than are more distal measures. Additional proximal outcomes that future research might explore include number of job offers received, as well as newcomer socialization and adjustment.

Conclusion

A number of researchers have called for research that demonstrates that applicant reactions really matter (e.g., Morgeson & Ryan, 2009; Ryan & Ployhart, 2000; Sackett & Lievens, 2008). The concern is that although applicant reactions have been found to relate to test scores and applicant intentions, the limited research that has examined relations between applicant reactions and more behavioral outcomes has yielded insignificant results (Hausknecht et al., 2004). In the current study, we developed a framework to help understand whether and when reactions may affect the critical outcome of job performance. Our findings indicated that certain types of test reactions can have both direct and indirect effects on job performance, particularly when they are dispositional in nature. In contrast, it does not appear that reactions affect (positively or negatively) the criterion-related validity of selection procedures for the prediction of job performance. We hope the framework we present will serve as a foundation for future work that seeks to further increase understanding of the effects of applicant reactions.

Footnotes

1 In line with this distinction, *mediated* effects are analogous to partial mediation, and *indirect* effects are analogous to full mediation (Mathieu & Taylor, 2006).
2 Two individuals were eliminated due to extreme scores on the criterion. Results of the analyses did not change as a result of removing these individuals.
3 We thank an anonymous reviewer for noting this possibility.
4 We were able to assess range restriction in Studies 1–3, and the resulting *u* values ranged from .86 to .97.

References

Aguinis, H., Culpepper, S. A., & Pierce, C. A. (2010). Revival of test bias research in preemployment testing. *Journal of Applied Psychology, 95*, 648–680. doi:10.1037/a0018714

Aguinis, H., & Stone-Romero, E. F. (1997). Methodological artifacts in moderated multiple regression and their effects on statistical power. *Journal of Applied Psychology, 82*, 192–206. doi:10.1037/0021-9010.82.1.192

Ajzen, I., & Fishbein, M. (1977). Attitude–behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84*, 888–918. doi:10.1037/0033-2909.84.5.888

Anderson, N., & Witvliet, C. (2008). Fairness reactions to personnel selection methods: An international comparison between the Netherlands, the United States, France, Spain, Portugal, and Singapore. *International Journal of Selection and Assessment, 16*, 1–13. doi:10.1111/j.1468-2389.2008.00404.x

Arvey, R. D., Strickland, W., Drauden, G., & Martin, C. (1990). Motivational components of test taking. *Personnel Psychology, 43*, 695–716. doi:10.1111/j.1744-6570.1990.tb00679.x

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*, 122–147. doi:10.1037/0003-066X.37.2.122

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1–26. doi:10.1111/j.1744-6570.1991.tb00688.x

Bartlett, C. J., Bobko, P., Mosier, S. B., & Hannan, R. (1978). Testing for fairness with a moderated multiple regression strategy: An alternative to differential analysis. *Personnel Psychology, 31*, 233–241. doi:10.1111/j.1744-6570.1978.tb00442.x

Bauer, T. N., Maertz, C. P., Jr., Dolen, M. R., & Campion, M. A. (1998). Longitudinal assessment of applicant reactions to employment testing and test outcome feedback. *Journal of Applied Psychology, 83*, 892–903. doi:10.1037/0021-9010.83.6.892

Bauer, T. N., & Truxillo, D. M. (2006). Applicant reactions to situational judgment tests. In J. A.Weekley & R. E.Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 233–249). Mahwah, NJ: Erlbaum.

Bauer, T. N., Truxillo, D. M., Sanchez, R. J., Craig, J. M., Ferrara, P., & Campion, M. A. (2001). Applicant reactions to selection: Development of the Selection Procedural Justice Scale (SPJS). *Personnel Psychology, 54*, 387–419. doi:10.1111/j.1744-6570.2001.tb00097.x

Bauer, T. N., Truxillo, D. M., Tucker, J. S., Weathers, V., Bertolino, M., Erdogan, B., & Campion, M. A. (2006). Selection in the information age: The impact of privacy concerns and computer experience on applicant reactions. *Journal of Management, 32*, 601–621. doi:10.1177/0149206306289829

Bonaccio, S., Reeve, C. L., & Winford, E. C. (2012). Text anxiety on cognitive ability test can result in differential predictive validity of academic performance. *Personality and Individual Differences, 52*, 497–502. doi:10.1016/j.paid.2011.11.015

Borman, W. C., White, L. A., Pulakos, E. D., & Oppler, S. H. (1991). Models of supervisory job performance ratings. *Journal of AppliedPsychology, 76*, 863–872. doi:10.1037/0021-9010.76.6.863

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N.Schmitt & W.Borman (Eds.), *Personnel selection in organizations* (pp. 35–71). San Francisco, CA: Jossey-Bass.

Chamorro-Premuzic, T., Ahmetoglu, G., & Furnham, A. (2008). Little more than personality: Dispositional determinants of test anxiety (the Big Five, core self-evaluations, and self-assessed intelligence). *Learning and Individual Differences, 18*, 258–263. doi:10.1016/j.lindif.2007.09.002

Chan, D., & Schmitt, N. (2004). An agenda for future research on applicant reactions to selection procedures: A construct-oriented approach. *International Journal of Selection and Assessment, 12*, 9–23. doi:10.1111/j.0965-075X.2004.00260.x

Chan, D., Schmitt, N., DeShon, R. P., Clause, C. S., & Delbridge, K. (1997). Reactions to cognitive ability tests: The relationships between race, test performance, face validity perceptions, and test-taking motivation. *Journal of Applied Psychology, 82*, 300–310. doi:10.1037/0021-9010.82.2.300

Chan, D., Schmitt, N., Jennings, D., Clause, C. S., & Delbridge, K. (1998). Applicant perceptions of test fairness integrating justice and self-serving bias perspectives. *International Journal of Selection and Assessment, 6*, 232–239. doi:10.1111/1468-2389.00094

Chan, D., Schmitt, N., Sacco, J. M., & DeShon, R. P. (1998). Understanding pretest and posttest reactions to cognitive ability and personality tests. *Journal of Applied Psychology, 83*, 471–485. doi:10.1037/0021-9010.83.3.471

Cleary, T. A. (1968). Test bias: Prediction of grades of Negro and White students in integrated colleges. *Journal of Educational Measurement, 5*, 115–124. doi:10.1111/j.1745-3984.1968.tb00613

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Routledge.

Cohen-Charash, Y., & Spector, P. E. (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes, 86*, 278–321. doi:10.1006/obhd.2001.2958

Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O. L. H., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology, 86*, 425–445. doi:10.1037/0021-9010.86.3.425

Cortina, J. M., & Folger, R. G. (1998). When is it acceptable to accept a null hypothesis: No way, Jose?*Organizational Research Methods, 1*, 334–350. doi:10.1177/109442819813004

Dye, D. A., Reck, M., & McDaniel, M. A. (1993). The validity of job knowledge measures. *International Journal of Selection and Assessment, 1*, 153–157. doi:10.1111/j.1468-2389.1993.tb00103.x

Eden, D., & Aviram, A. (1993). Self-efficacy training to speed reemployment: Helping people to help themselves. *Journal of Applied Psychology, 78*, 352–360. doi:10.1037/0021-9010.78.3.352

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & Emotion, 6*, 409–434. doi:10.1080/02699939208409696

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*, 336–353. doi:10.1037/1528-3542.7.2.336

Ford, D. K., Truxillo, D. M., & Bauer, T. N. (2009). Rejected but still there: Shifting the focus in applicant reactions to the promotional context. *International Journal of Selection and Assessment, 17*, 402–416. doi:10.1111/j.1468-2389.2009.00482.x

Ford, M. T., Cerasoli, C. P., Higgins, J. A., & Decesare, A. L. (2011). Relationships between psychological, physical, and behavioural health and work performance: A review and meta-analysis. *Work & Stress, 25*, 185–204. doi:10.1080/02678373.2011.609035

Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review, 18*, 694–734. doi:10.2307/258595

Gilliland, S. W. (1994). Effects of procedural and distributive justice on reactions to a selection system. *Journal of Applied Psychology, 79*, 691–701. doi:10.1037/0021-9010.79.5.691

Harrison, D. A., Newman, D. A., & Roth, P. L. (2006). How important are job attitudes? Meta-analytic comparisons of integrative behavioral outcomes and time sequences. *Academy of Management Journal, 49*, 305–325. doi:10.5465/AMJ.2006.20786077

Hattrup, K., Schmitt, N., & Landis, R. S. (1992). Equivalence of constructs measured by job-specific and commercially available aptitude tests. *Journal of Applied Psychology, 77*, 298–308. doi:10.1037/0021-9010.77.3.298

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. doi:10.1111/j.1744-6570.2004.00003.x

Hülsheger, U. R., & Anderson, N. (2009). Applicant perspectives in selection: Going beyond preference reactions. *International Journal of Selection and Assessment, 17*, 335–345. doi:10.1111/j.1468-2389.2009.00477.x

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*, 72–98. doi:10.1037/0033-2909.96.1.72

Judge, T. A., & Bono, J. E. (2001). Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology, 86*, 80–92. doi:10.1037/0021-9010.86.1.80

Krausz, M., Koslowsky, M., & Eiser, A. (1998). Distal and proximal influences on turnover intentions and satisfaction: Support for a withdrawal progression theory. *Journal of Vocational Behavior, 52*, 59–71. doi:10.1006/jvbe.1996.1565

Lievens, F., Buyse, T., & Sackett, P. R. (2005). The operational validity of a video-based situational judgment test for medical college admissions: Illustrating the importance of matching predictor and criterion construct domains. *Journal of Applied Psychology, 90*, 442–452. doi:10.1037/0021-9010.90.3.442

Macan, T. H., Avedon, M. J., Paese, M., & Smith, D. E. (1994). The effects of applicants' reactions to cognitive ability tests and an assessment center. *Personnel Psychology, 47*, 715–738. doi:10.1111/j.1744-6570.1994.tb01573.x

Masterson, S. S., Lewis, K., Goldman, B. M., & Taylor, M. S. (2000). Integrating justice and social exchange: The differing effects of fair procedures and treatment on work relationships. *Academy of Management Journal, 43*, 738–748. doi:10.2307/1556364

Mathieu, J. E., & Taylor, S. R. (2006). Clarifying conditions and decision points for mediational type inferences in organizational behavior. *Journal of Organizational Behavior, 27*, 1031–1056. doi:10.1002/job.406

McCarthy, J., & Goffin, R. (2004). Measuring job interview anxiety: Beyond weak knees and sweaty palms. *Personnel Psychology, 57*, 607–637. doi:10.1111/j.1744-6570.2004.00002.x

McCarthy, J., Hrabluik, C., & Jelley, R. B. (2009). Progression through the ranks: Assessing employee reactions to high-stakes employment testing. *Personnel Psychology, 62*, 793–832. doi:10.1111/j.1744-6570.2009.01158.x

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

McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology, 86*, 730–740. doi:10.1037/0021-9010.86.4.730

Meade, A. W., & Tonidandel, S. (2010). Not seeing clearly with Cleary: What test bias analyses do and do not tell us. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 3*, 192–205. doi:10.1111/j.1754-9434.2010.01223.x

Morgeson, F. P., & Ryan, A. M. (2009). Reacting to applicant perspectives research: What's next?*International Journal of Selection and Assessment, 17*, 431–437. doi:10.1111/j.1468-2389.2009.00484.x

Murphy, K. R., & Shiarella, A. H. (1997). Implications of the multidimensional nature of job performance for the validity of selection tests: Multivariate frameworks for studying test validity. *Personnel Psychology, 50,* 823–854. doi:10.1111/j.1744-6570.1997.tb01484.x

O'Neill, T. A., Goffin, R. D., & Gellatly, I. R. (2010). Test-taking motivation and personality test validity. *Journal of Personnel Psychology, 9,* 117–125. doi:10.1027/1866-5888/a000012

Petty, R., & Cacioppo, J. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York, NY: Springer-Verlag.

Poppen, R. (1988). *Behavioral relaxation training and assessment*. Elmsford, NY: Pergamon Press.

Posthuma, R. A., & Campion, M. A. (2005). When do multiple dimensions of procedural justice predict agreement to publicly endorse your employer in recruitment advertisements?*Journal of Occupational and Organizational Psychology, 78,* 431–452. doi:10.1348/096317905X27173

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments & Computers, 36,* 717–731. doi:10.3758/BF03206553

Proost, K., Derous, E., Schreurs, B., Hagtvet, K. A., & De Witte, K. (2008). Selection test anxiety: Investigating applicants' self- vs. other-referenced anxiety in a real selection setting. *International Journal of Selection and Assessment, 16,* 14–26. doi:10.1111/j.1468-2389.2008.00405.x

Reeve, C. L., Heggestad, E. D., & Lievens, F. (2009). Modeling the impact of test anxiety and test familiarity on the criterion-related validity of cognitive ability tests. *Intelligence, 37,* 34–41. doi:10.1016/j.intell.2008.05.003

Roth, P. L., Bobko, P., & McFarland, L. A. (2005). A meta-analysis of work sample test validity: Updating and integrating some classic literature. *Personnel Psychology, 58,* 1009–1037. doi:10.1111/j.1744-6570.2005.00714.x

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

Ryan, A. M., & Ployhart, R. E. (2000). Applicants' perceptions of selection procedures and decisions: A critical review and agenda for the future. *Journal of Management, 26,* 565–606. doi:10.1177/014920630002600308

Sackett, P. R., & Lievens, F. (2008). Personnel selection. *Annual Review of Psychology, 59,* 419–450. doi:10.1146/annurev.psych.59.103006.093716

Salgado, J. F., Remeseiro, C., & Iglesias, M. (1996). Personality and test-taking motivation. *Psicothema, 8,* 553–562.

Sarason, I. G., Sarason, B. R., & Pierce, G. R. (1990). Anxiety, cognitive interference, and performance. *Journal of Social Behavior and Personality, 5,* 1–18.

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. doi:10.1037/a0018920

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124,* 262–274. doi:10.1037/0033-2909.124.2.262

Schmit, M. J., & Ryan, A. M. (1992). Test-taking dispositions: A missing link?*Journal of Applied Psychology, 77,* 629–637. doi:10.1037/0021-9010.77.5.629

Schmitt, N., & Chan, D. (2006). Situational judgment tests: Method or construct? In J. A.Weekley & R. E.Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 135–155). Mahwah, NJ: Erlbaum.

Schmitt, N., Cortina, J. M., Ingerick, M. J., & Wiechmann, D. (2003). Personnel selection and employee performance. In *Handbook of psychology*(pp. 77–105). Hoboken, NJ: Wiley. doi:10.1002/0471264385.wei1205

Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, *37*, 407–422. doi:10.1111/j.1744-6570.1984.tb00519.x

Schneider, R. J., Hough, L. M., & Dunnette, M. D. (1996). Broadsided by broad traits: How to sink science in five dimensions or less. *Journal of Organizational Behavior*, *17*, 639–655.

Smither, J. W., Reilly, R. R., Millsap, R. E., & Pearlman, K. (1993). Applicant reactions to selection procedures. *Personnel Psychology*, *46*, 49–76. doi:10.1111/j.1744-6570.1993.tb00867.x

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, *44*, 703–742. doi:10.1111/j.1744-6570.1991.tb00696.x

Thorsteinson, T. J., & Ryan, A. M. (1997). The effect of selection ratio on perceptions of the fairness of a selection test battery. *International Journalof Selection and Assessment*, *5*, 159–168. doi:10.1111/1468-2389.00056

Truxillo, D. M., & Bauer, T. N. (2010). Applicant reactions to organizations and selection systems. In S.Zedeck (Ed.), APA handbook of industrial and organizational psychology: Vol. *2*. *Selecting and developing members for the organization* (pp. 379–397). Washington, DC: American Psychological Association. doi:10.1037/12170-012

Truxillo, D. M., Bauer, T. N., Campion, M. A., & Paronto, M. E. (2002). Selection fairness information and applicant reactions: A longitudinal field study. *Journal of Applied Psychology*, *87*, 1020–1031. doi:10.1037/0021-9010.87.6.1020

Weiss, H. M., & Cropanzano, R. (1996). Affective events theory: A theoretical discussion of the structure, causes, and consequences of affective experiences at work. *Research in Organizational Behavior*, *18*, 1–74.

Wernimont, P. F., & Campbell, J. P. (1968). Signs, samples, and criteria. *Journal of Applied Psychology*, *52*, 372–376. doi:10.1037/h0026244

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, *37*, 197–206. doi:10.1086/651257