Racial differences in promotion candidate performance and reactions to selection procedures: a field study in a diverse top-management context

JOHN BRET BECTON1*, HUBERT S. FEILD2, WILLIAM F. GILES2 AND ALLISON JONES-FARMER2

1Winthrop University, Management and Marketing Department, Rock Hill, SC, U.S.A.
2Auburn University, Department of Management, Auburn, AL, U.S.A.

Summary
The present study examined African-American and White promotion candidates’ reactions to and performance on selection procedures that were completed within a police department where African Americans occupied the majority of top-management positions. Reactions (perceived job relatedness and test-taking motivation) of 187 candidates competing for promotion to the rank of sergeant were assessed after completing a written job knowledge test and a situational interview. Analyses showed that both the African-American and White candidates judged the situational interview to be more job-related than the pencil-and-paper job knowledge test. In addition, African Americans perceived both selection measures to be more job-related and reported higher levels of test-taking motivation than White candidates even though African Americans performed more poorly than White candidates on the paper-and-pencil test. These results challenge the contention that lower test-taking motivation for African-American candidates is related to racial differences in performance on pencil-and-paper tests. Implications and directions for future research on reactions to selection procedures for promotion in racially diverse employment settings are discussed. Copyright © 2007 John Wiley & Sons, Ltd.

Introduction
Over the past decade, few topics have received as much attention in the industrial psychology/human resources management literature as employment testing. A large volume of research on employment testing has focused on the validity and utility of employee selection devices (see Hunter & Hunter, 1984; Hunter, Schmidt, & Coggin, 1988; Salgado, Viswesvaran, & Ones, 2001; Schmidt, 1988; Schmidt & Hunter, 1998). However, in recent years, researchers have increased efforts to investigate...
candidate reactions to selection tests (Anderson, Born, & Cunningham-Snell, 2001). Research on test reactions has typically focused on characteristics such as fairness, job relatedness, test-taking motivation, and other test-taking attitudes. Research has examined candidate reactions to selection measures such as pencil-and-paper tests, structured interviews, biographical data, personality inventories, assessment centers, and work samples (e.g., Bauer, Truxillo, Paronto, Weekley, & Campion, 2004; Chan & Schmitt, 1997; Chan, Schmitt, Deshon, Clause, & Delbridge, 1997; Hausknecht, Day, & Thomas, 2004; Macan, Avedon, Paese, & Smith, 1994; Schmitt, Oswald, Kim, Gillespie, & Ramsay, 2004; Truxillo, Bauer, Campion, & Paronto, 2002).

The extant literature suggests that candidates prefer methods having greater psychological and physical fidelity with work settings, such as work simulations, as opposed to methods having less fidelity, such as pencil-and-paper tests. Selection procedures that simulate actual job behaviors such as work samples, in-baskets, and role-plays have been viewed as having more face validity and perceived more favorably than pencil-and-paper methods (Dodd, 1977; Macan et al., 1994; Schmidt, Greenthal, Hunter, Berner, & Seaton, 1977; Smither, Reilly, Millsap, Pearlman, & Stoffey, 1993; Smither & Pearlman, 1991; Steiner & Gilliland, 1996). Furthermore, some researchers suggest that African Americans tend to view simulations and verbally administered tests more positively than pencil-and-paper tests in comparison with Whites (Chan & Schmitt, 1997; Goldstein, Braverman, & Chung, 1998; Helms, 1992; Ryan & Greguras, 1997) although little research has tested this assumption.

**Rationale for the present study**

While results of the various studies cited above are important, there are three reasons why additional study of applicant test reactions is necessary. First, the role of organizational context factors such as the presence of minority leadership on applicants’ reactions has not been explored (Ryan & Ployhart, 2000). Our review of the literature revealed that no study has explicitly examined racial differences in applicant reactions in organizations where top management is predominantly African-American. Although data concerning the demographic composition of organizations’ top management has not generally been reported in previous studies of reactions to testing, many researchers have addressed the fact that African Americans continue to be dramatically under-represented in leadership and managerial positions across the U.S. (DiTomaso & Thompson, 1988; Dovidio & Gaertner, 1996; This Nation, 2001). According to the Equal Employment Opportunity Commission (EEOC), fewer than 10 per cent of officials and managers are African-American, with over 80 per cent being White (EEOC, 2003). Since African Americans presently comprise less than 10 per cent of top-management positions (EEOC, 2003), we can likely assume that few, if any, of the published studies included organizations where the majority of senior leadership was African-American. The present study examined promotion candidate reactions to pencil-and-paper tests and situational interviews in an organization with predominant African-American leadership.

Second, most test reaction studies have been conducted in lab settings using either undergraduate students or individuals working in simulated employment settings. The stakes are certainly higher for actual job candidates, and one might expect the reactions and motivation of actual job candidates to differ from those in lab studies (Arvey, Strickland, Drauden, & Martin, 1990). Thus, it is unclear if the findings of much existing research can be generalized to actual selection contexts. In this study, our participants were actual employees competing for promotion in a real organization.

Third, the vast majority of test reaction investigations have been conducted in an entry-level context (i.e., external applicants attempting to gain entry into the organization). While test reactions of entry-level applicants are certainly important, reactions of promotion candidates are perhaps even more
critical, yet their reactions to various assessment devices have been largely neglected in the literature. Understanding promotion candidate reactions is important because incumbents’ reactions are felt throughout the organization, producing negative consequences that can be more acutely felt by the organization than those in the entry-level context (Truxillo & Bauer, 1999). In the present study, we compared the reactions of promotion candidates to two selection procedures (a situational interview versus a pencil-and-paper job knowledge test) used in promotion decision-making. Specifically, we examined two forms of reactions: (a) candidates’ perceptions of how well the selection procedures were related to the job and (b) candidates’ motivation to perform well on these procedures.

**Job-relatedness**

Job-relatedness refers to the extent to which a selection procedure either appears to measure content relevant to the job situation or appears to be valid (Bauer, Truxillo, Sanchez, Craig, Ferra, & Campion, 2001). Job-relatedness is important to study because it may have the greatest impact on fairness perceptions as compared to other formal characteristics of selection procedures (Gilliland, 1993). Face validity is another popular topic in the reactions-to-testing literature. Face validity is the degree to which candidates perceive the content of a selection device to be related to the content of the job (Smither et al., 1993). There is overlap in these definitions, and a clear distinction between the concepts as they relate to test reactions is difficult to draw. While some researchers have made a distinction between perceived job-relatedness and face validity (cf. Elkins & Phillips, 2000), other researchers have argued that they are similar, highly related concepts (Chan & Schmitt, 2004). We view face validity and perceived job-relatedness as highly similar, if not interchangeable, concepts from the candidates’ perspective, and we draw on research using both concepts to derive our hypotheses.

Research concerning perceptions of job-relatedness has yielded some important findings. Studies have shown that selection procedures that simulate actual job behaviors such as situational interviews, work samples, in-baskets, and role-plays are viewed as having more face validity, and are perceived more favorably than pencil-and-paper methods (Macan, Avedon, Paese, & Smith, 1994; Schmidt et al., 1977; Smither et al., 1993; Smither & Pearlman, 1991). Chan and Schmitt (1997) found that the high degree of face validity and positive candidate reactions to simulations are frequently attributed to their realistic test situation and similarity to the target job (i.e., high fidelity). Furthermore, job-relatedness has been shown to be highly correlated with perceived fairness (Schmitt, Gilliland, Landis, & Devine, 1993), such that reactions to selection tests are substantially improved by framing test content around business-related topics and increasing the items’ contextual relevance (Rynes & Connerley, 1993). Moreover, job applicants prefer tests with obviously job-related content (Rynes & Connerley, 1993). In general, research has shown that selection devices such as interviews and work samples are perceived more favorably than cognitive ability tests (Hausknecht et al., 2004). Situational interviews are viewed as more job related than paper-and-pencil tests because situational questions pose hypothetical situations that are likely to occur on the job and require candidates to explain how they would handle these situations. As a result, situational interviews have a high degree of face validity, and candidates can more easily determine that the questions are related to the demands or duties of the job. In contrast, pencil-and-paper tests are often devoid of such organizational or situational context and candidates consequently have a more difficult time recognizing the linkages between the question content and content of the job. Because prior research on job-relatedness perceptions generally indicates that interviews are viewed as more job-related than more traditional testing methods (i.e., pencil-and-paper), we advance the following hypothesis:

**Hypothesis 1**: Promotion candidates will perceive a situational interview as having greater job-relatedness than a pencil-and-paper test.
Test-taking motivation
Test-taking motivation refers to the extent to which applicants try their best to perform well on a test and view the results of the test as important (Arvey et al., 1990). Much of the research involving test-taking motivation has centered on the relationship between applicants’ motivation and test performance. For example, several studies have concluded that applicants’ test-taking motivation is associated with test performance (e.g., Arvey et al., 1990; Chan et al., 1997); however, other investigators have not found such a relationship (e.g., Sanchez, Truxillo, & Bauer, 2000; Schmit & Ryan, 1997). Moreover, negative reactions to selection procedures are thought to be linked to reduced motivation to do well on such procedures (Arvey et al., 1990). For example, the face validity of a selection device seems to influence applicant motivation to do well on the test (Robertson & Kandola, 1982). In other words, the more transparent the linkage between the content of the test and the content of the job, the greater the motivation to do well on the test. Based on the previous hypothesis that situational interviews will be perceived as being more job-related in comparison with pencil-and-paper tests and the fact that job-relatedness and test-taking motivation appear to be related (Robertson & Kandola, 1982), it seems reasonable to propose that situational interviews would be associated with higher test-taking motivation. Therefore, we hypothesize that

**Hypothesis 2:** Promotion candidates will report higher test-taking motivation when completing a situational interview than when taking a pencil-and-paper test.

Racial differences in reactions to and performance on selection procedures

Within the reactions-to-testing literature, one area that has received relatively little attention is whether there are racial differences in reactions to assessment and whether such differences exist for various assessment methods, such as pencil-and-paper tests or employment interviews. While African-American/White differences in test performance have generated ample research, racial differences in reactions to selection procedures have not received the same level of attention in the literature. Racial differences in reactions to selection procedures are clearly important as they could have implications for minority recruitment programs, adverse impact determination, and equal employment litigation. Moreover, African-Americans’ performance on cognitive tests may be negatively influenced by their attitudes toward testing, and performance on these types of tests might be improved if these attitudes can be modified (McKay & Doverspike, 2001). Arvey, Strickland, Drauden, and Martin (1990) examined African-American/White differences in motivation toward employment tests and found that lack of test-taking motivation among African-American candidates undermined their subsequent performance on employment tests. Other studies have indicated that racial differences in the perceived fairness of cognitive ability tests may explain or be explained by racial differences in test-taking motivation and performance on tests (Chan, Schmitt, Deshon, Clause, & Delbridge, 1997; Chan, Schmitt, Jennings, Clause, & Delbridge, 1998; Chan, Schmitt, Sacco, & DeShon, 1998). Additionally, some researchers have suggested that the racial gap in cognitive ability test performance may be due in part to African Americans having less favorable test-taking motivation and test attitudes (e.g., Arvey et al., 1990; Chan et al., 1997; Helms, 1992). Consequently, the study of racial differences in test reactions and test-taking motivation is an important element to understanding the fairness perceptions of promotion candidates to the assessment process within an organization.

In general, African-Americans’ test reactions are more negative than those of Whites, but racial differences are smaller for interactive types of tests than for pencil-and-paper tests (Chan & Schmitt, 2004). However, empirical evidence concerning racial differences in perceptions of job-relatedness is meager (Chan et al., 1997), and previous studies of racial differences in job-relatedness perceptions...
have produced mixed results. Schmit and Ryan (1997) and Smither et al. (1993), for example, reported no differences in face validity perceptions between African Americans and Whites. Apparently, only two other studies have examined racial differences in the perceived job-relatedness of tests. Chan et al. (1997) found lower levels of face validity perceptions for African Americans after completing a cognitive ability test battery. In addition, Chan and Schmitt (1997) identified a race × method interaction effect such that the differences in face validity perceptions of African Americans and Whites were greater for a pencil-and-paper test than for a video-based method of testing.

As for racial differences in test-taking motivation, several researchers have theorized that African Americans tend to have somewhat negative attitudes concerning employment testing. Frierson (1986) argued that African Americans have less opportunity to learn and practice the skills necessary to do well on standardized tests, thereby reducing their belief in the tests. Other studies have noted African-American test-taking motivation to be significantly lower than that of Whites (Arvey et al., 1990; Arvey et al., 1991; Chan et al., 1998). Likewise, Chan et al. (1997) reported lower levels of test-taking motivation for African Americans after completing a cognitive ability test battery. Some researchers have suggested that minority applicants view performance assessments more favorably than multiple-choice testing (Ryan & Greguras, 1998) although little empirical research has tested this assumption.

To summarize, there is some evidence that African Americans have lower levels of test-taking motivation than Whites and inconclusive evidence concerning racial differences in job-relatedness perceptions. What is not clear from prior research, however, is the role that organizational context plays in applicant reactions to testing (Arvey & Sackett, 1993). It is likely that many of the prior studies in this area were conducted in organizations having predominantly White leadership. Because our study was conducted in an organization with predominantly African-American leadership, it is important to balance our understanding of prior research with the emerging dynamics of an organization having a different racial demographic.

**Relationship between organizational context and reactions to testing**

While factors such as race, test type, construct measured, and test preparation have been thoroughly studied, many other factors likely contribute to the formation of reactions to tests. Researchers have suggested that organizational history, selection ratio, and organizational resources may affect reactions to testing (Arvey & Sackett, 1993) as well. In civil service settings, the public nature of testing, presence of strong affirmative action efforts, and past discrimination seem to have important effects on applicant perceptions of selection processes (Truxillo & Bauer, 1999). However, a systematic examination of the role of these contextual factors has not been conducted (Arvey & Sackett, 1993). Furthermore, Ryan and Ployhart (2000) noted that an accumulation of research from different contexts might be required before any conclusions concerning the influence of organizational context can be made, and ignoring its role in the reactions to testing literature may result in potentially misleading conclusions.

It has been suggested that organizational context influences candidate reactions (Arvey & Sackett, 1993), and that racial differences in real-world settings are highly influenced by context (Ryan & Ployhart, 2000). Milliken and Martins (1996) hypothesized that racial diversity in organizations, particularly diversity at the supervisory level has important implications for the behavior and perceptions of minority employees. Research findings suggest that similarity between an employee and members of their work unit and/or supervisor is associated with organizational commitment, turnover...
intentions, absenteeism, career satisfaction, and perceptions of organizational acceptance and discretion (Greenhaus, Parasuraman & Wormley, 1990; Tsui, Egan, & O’Reilly, 1992). Additionally, Ely (1994) suggested that diversity in the composition of organizational groups, and in top management specifically, may be important because employees behave differently when they perceive that they have access to power and opportunity compared to when they perceive the organization as less supportive of their development.

Furthermore, research in organizations with a history of discriminatory practices and/or the presence of a strong affirmative action program has indicated that these contextual factors influence perceptions of fairness (Ryan, Ployhart, Greguras, & Schmit, 1997; Ryan, Sacco, McFarland, & Kriska, 2000; Schmit & Ryan, 1997; Truxillo & Bauer, 1999). For example, African Americans have more positive views than Whites do about the fairness of testing in organizational settings where there are strong affirmative action programs and minorities in highly visible leadership positions (Ryan, Ployhart, Greguras, & Schmit, 1997; Schmit & Ryan, 1997). Therefore, it would seem that an organizational context which reflects diversity, equal opportunity, and the absence of discrimination would have a positive influence on the test perceptions of African Americans. In this context, African-American candidates likely view the prevalence of African Americans in leadership positions as an indication of the organization’s commitment to fairness and equal opportunity. As a result, we expect that the perceptions (i.e., job-relatedness and test-taking motivation) of African Americans in this study will be more positive than the perceptions of Whites. The following hypotheses are offered:

**Hypothesis 3a:** In an organization where the majority of leadership positions are occupied by African Americans, African-American promotion candidates’ perceptions of the job-relatedness of selection procedures will be more positive than those of White candidates.

**Hypothesis 3b:** In an organization where the majority of leadership positions are occupied by African Americans, African-American promotion candidates’ test-taking motivation will be higher than that of White candidates.

Less clear is how racially diverse top management affects the interaction between race of test taker and test type on test-taking motivation and job-relatedness. Evidence suggests that African Americans, in general, prefer work samples, simulations, and orally administered selection devices (Helms, 1992; Goldstein, Braverman, & Chung, 1993). For example, Chan and Schmitt (2004) found smaller racial differences in reactions to interactive tests compared to pencil-and-paper tests. Some have hypothesized that the reason for this difference lies in the dense written content of pencil-and-paper tests. Helms (1992) argued that tests administered orally may influence the test responses of African-American candidates, which may explain the absence of African-American–White differences in structured interviews. Citing Helms (1992), Goldstein, et al. (1992) theorized that African-American values and beliefs stress communalism, movement, and orality, which would consequently impact their test performance such that African Americans perform better on orally administered tests. Furthermore, Frierson (1986) reasoned that African Americans have less opportunity to learn and practice the skills necessary to do well on written standardized tests, thereby reducing their belief in these tests. Other studies have found African-American test-taking motivation to be significantly lower than that of Whites (Arvey et al., 1990; Chan, Schmitt, DeShon, Clause, & Delbridge, 1997). Some researchers have suggested that minority applicants view performance assessments more favorably than multiple-choice testing (Ryan & Greguras, 1998) although little research has been conducted on this assumption.

As mentioned earlier, very few studies have examined racial differences in the perceived job-relatedness of tests. Chan et al. (1997) found face validity perceptions and test-taking motivation
for African Americans to be lower than those of Whites after completing a cognitive ability test battery. Additionally, Chan and Schmitt (1997) reported greater differences in face validity perceptions of African Americans and Whites for a pencil-and-paper test when compared to a video-based test. Smither et al. (1993) studied group differences in reactions to selection procedures, including perceived face validity. However, the authors reported racial differences in perceived knowledge of results and likelihood for improvement but failed to mention racial differences in perceived face validity.

On the surface, there appears to be ample evidence to expect that African-American candidates will view situational interviews more favorably than pencil-and-paper tests due to the interviews’ absence of written content and their aural nature. However, this expectation ignores the role of organizational context in the formation of test reactions. Could the presence of highly visible minority leadership reduce racial differences in reactions to different methods of testing? Previous research on the similar-to-me effect suggests that individuals prefer people and situations where they believe people and situations are similar to themselves (Clark & Fiske, 1982). Perceived similarity between people has been shown to determine likeability and trust in general psychological research (Patzer, 1985) and in employment settings in particular (Kanter, 1977). Similarly, social identification theory suggests that individuals’ self-identities are defined partly by their membership in groups that they especially value and find emotionally significant (Tajfel, 1982). When perceiving and evaluating others, individuals are likely to judge more favorably those individuals who hold common group membership, such as a racial group (Tajfel, 1982). Both the similar-to-me effect and social identification theory suggest that members of a particular group (e.g., racial group) will perceive or evaluate members of the same group more favorably than members of other groups (Prewett-Livingston, Feild, Veres, & Lewis, 1996). It may be possible that a similar type of relationship will exist between the racial composition of an organization’s top leadership and minority applicants’ perceptions of test fairness. In other words, in a context of predominantly African-American leadership, African Americans may recognize and value the fact that they are in the same racial group as the top leaders and decision makers of the organization, which may have strong, important effects on their fairness perceptions. Therefore, in such a context, perhaps African Americans base their perceptions of the selection process on these factors more so than the characteristics of the tests themselves, resulting in positive reactions to the tests regardless of test type. Therefore, we formed the following research question:

**Research Question 1**: In an organization where the majority of leadership positions are occupied by African Americans, will the interaction of promotion candidates’ race and the type of selection procedure they complete be associated with how they view these selection procedures (in terms of perceived job-relatedness and test-taking motivation)?

Some researchers suggest that a partial explanation for the racial gap in test performance seen with pencil-and-paper tests is due to racial differences in test perceptions and test-taking motivation (Arvey et al., 1990; Chan & Schmitt, 1997; Steele & Aronson, 1995). In essence, these researchers have suggested that African Americans generally have more negative perceptions of pencil-and-paper tests. This likely has a negative influence on their test-taking motivation, which, in turn, undermines subsequent test performance. If the presence of African Americans in top-management positions within the organization results in more positive perceptions of tests and test-taking motivation, regardless of test type, one might expect to see race-related performance differences ameliorated. As a result, we pose the following research question:

**Research Question 2**: In an organization where the majority of leadership positions are occupied by African Americans, will the interaction of promotion candidates’ race and the type of selection procedures they complete be associated with candidates’ performance on the selection procedures?
Organization Context

Data for this study were collected from the Dalton County Police Department (a pseudonym) located in a metropolitan area in the southeastern United States. The police department has approximately 1,000 sworn officers consisting of approximately 47% African Americans, 47% Whites, and 6% of other racial/ethnic categories.

Societal Context
Dalton County has a population of approximately 700,000 residents that is culturally diverse including over 64 spoken languages and a melting pot of well-educated communities. With a population that is 44% White and 44% African-American and a remaining mixture of Hispanic and Asian residents, Dalton County is the most culturally diverse county in the Southeast. The per capita income for the county is about $24000.

Historical Context
As discussed above, Dalton County is a very diverse area. However, neither the county nor the police department has been as diverse as it is today. Nearly two decades ago, the power structure within the police department and other county offices was primarily White. As the demographics in Dalton County became more diverse, there were numerous complaints, accusations, and litigation concerning the unfairness of hiring and promotion practices of county departments, including the police department, during this time. As a result, the police department took numerous measures to increase the diversity among its officers and leadership including an affirmative action program and hiring a consulting firm to develop fair and content-valid selection procedures.

Method

Organizational setting and participants
Data for this study were collected from candidates competing for promotion to the rank of sergeant in a large, metropolitan police department in the southeastern U.S. The police department was racially diverse, with over half (54%) of the candidates being non-White. Additionally, the department’s leadership included numerous minorities, especially in the upper-levels of management. African Americans occupied approximately 74% of the leadership positions within the organization (i.e., top and middle management positions), including the current police chief, the previous police chief, and commanders of most of the departmental units. Furthermore, since selection and promotion procedures are often associated with an organization’s human resources department, perceived fairness of procedures should be related to perceptions of the human resources department (Truxillo, Steiner, & Gilliland, 2004). The human resources department in which this study occurred was comprised of mostly African-American employees, including the human resources director. However, this organization has not always been as diverse as it is today. Approximately 15 years ago, the leadership within the police and human resources departments was primarily White. There were numerous
complaints, accusations, and litigation concerning the unfairness of hiring and promotion practices during this time. As a result, the department hired a consulting firm to develop fair and content-valid selection procedures and has continued to use these services for the past 10 years.

As part of an assessment center, a multiple-choice, job knowledge exam and situational interview were administered to all candidates \( (N = 203) \), and performance in these two components determined whether candidates progressed to the final stage of the selection procedures. Since the study examined differences in perceptions of African-American and White candidates, 12 participants of other racial groups were eliminated from the study. The remaining 191 candidates for promotion were solicited for participation in the study. All (100%) of the 191 candidates agreed to participate; however, four of the participant records contained missing information. Thus, 187 complete cases were available for analysis. Of these, 94 (51%) were African-American, and 93 (49%) were White. The sample consisted of 173 males (93%), and the majority (73%) had over 26 years of tenure with the organization.

**Measures**

**Job-relatedness**
Perceptions of *job-relatedness* for both the pencil-and-paper test and situational interview were measured with three items from Bauer, et al. Campion’s (2001) Selection Procedural Justice Scale (SPJS). Items include “Doing well on this test means a person can do the Sergeant job well,” “A person who scored well on this test will be a good Sergeant,” and “The actual content of the test was clearly related to the job of Sergeant.” Item responses were made on a response scale ranging from 1 = strongly disagree to 5 = strongly agree. Coefficient alphas for the job-relatedness ratings for both the pencil-and-paper test and situational interview were 0.79.

**Test-taking motivation**
Applicant *test-taking motivation* was measured by three items from the Test Attitude Survey (Arvey et al., 1990). Items composing the scale are “Doing well on this test was important to me,” “While taking this test, I concentrated and tried to do well,” and “I pushed myself to do well on this test.” Responses were made on a response scale ranging from 1 = strongly disagree to 5 = strongly agree. Coefficient alphas for the test-taking motivation measure for the pencil-and-paper test and situational interview were 0.77 and 0.73, respectively. While the alphas of the test-taking motivation and job-relatedness measures were not high, they are within an acceptable range, *i.e.* \( \alpha > 0.70 \) (Nunnally, 1978).

**Pencil-and-paper test performance**
The pencil-and-paper test consisted of 140 multiple-choice items measuring technical and departmental knowledge (*i.e.* department policies and procedures, laws, regulations, and law enforcement techniques). Each multiple-choice question had four response options with only one correct answer. Participants’ test score performance was based on the percentage of total correct items. The split-half reliability estimate (Spearman-Brown corrected) for the pencil-and-paper test score was 0.85. The candidates’ test scores were standardized to obtain values with a mean of zero and a standard deviation of one.

**Situational interview performance**
The situational interview consisted of questions developed to assess five performance dimensions. Based on critical incidents obtained from subject matter experts, hypothetical scenarios were developed that described actual situations a sergeant could be expected to encounter on the job. Each
scenario presented candidates with a hypothetical situation in which they were required to assume a supervisory role and describe the actions they would take to arrive at a resolution. The situational interview questions were level 3–4 (moderate to full constraints on both questions and scoring) according to Huffcutt and Arthur’s (Huffcutt & Arthur, 1994) interview structure typology.

The situational interviews consisted of four situational questions measuring technical and departmental knowledge, human relations, problem analysis, management ability, and oral communication, and each dimension was weighted according to the relative importance of the knowledge, skills and abilities (KSAs) comprising the dimensions as determined by a job analysis. Candidates were randomly assigned to one of three, two-person panels of trained interviewers for scoring. An overall score for the situational interview was computed by multiplying the average dimension score by the dimension weights and then summing the products. The inter-rater reliability (intraclass correlation) for the situational interview was 0.70. The candidates’ situational interview scores were standardized to obtain values with a mean of zero and a standard deviation of one.

Demographic data
Race, gender, and organizational tenure were collected via a self-report questionnaire. Race was coded “0” for African-American and “1” for White. Gender was coded “0” for male and “1” for female. We collected organizational tenure by asking respondents to use a 9-point rating scale (1 = 2–4 years; 9 = 26 or more years) to report their approximate tenure with the organization.

Procedures

The pencil-and-paper test and situational interview were implemented in an assessment center developed for the police department in which the study took place. The assessment center selection exercises were part of a larger project designed to create content-valid promotional procedures for the position of sergeant. Content and scoring for the selection measures were based on a thorough job analysis of the sergeant’s position. Candidates first completed the pencil-and-paper test and then completed the situational interview. After completing both test components, candidates completed surveys collecting their reactions to both tests. Candidates were told that participation in the study was voluntary; they would not receive any special benefit for participating or punishment for not participating in the study.

Since procedural and outcome fairness are related (Brockner & Wisenfeld, 1996), measuring procedural justice (i.e., perceived job-relatedness) after the results are known (i.e., test/interview scores are released and hiring decisions are made) may create a potential confound. Therefore, immediately after completing each of the selection components and prior to receiving their scores, candidates completed a survey to assess their reactions to the selection component just finished. The survey was described as part of a university research project not affiliated with the police department or the human resources department.

Results

Bivariate intercorrelations and reliability estimates for all study variables are presented in Table 1. Means and standard deviations for the variables are shown in Table 2. We assessed the structure of the job-relatedness and test-taking motivation measures using confirmatory factor analysis (CFA). The CFA resulted in an adjusted chi-square ($\chi^2/df$) of 2.45, a goodness of fit index (GFI) of 0.98, a
<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Gender (0 = male, 1 = female)</td>
<td>0.91</td>
<td>0.28</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Organizational tenure</td>
<td>3.48</td>
<td>1.62</td>
<td>0.05</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent variable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Race (0 = African American, 1 = White)</td>
<td>0.51</td>
<td>0.50</td>
<td>0.04</td>
<td>−0.19*</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dependent variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pencil-and-paper test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Job relatedness</td>
<td>2.49</td>
<td>0.85</td>
<td>−0.10</td>
<td>−0.03</td>
<td>0.19**</td>
<td>(0.79)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Test-taking motivation</td>
<td>4.43</td>
<td>0.56</td>
<td>0.10</td>
<td>0.01</td>
<td>0.22**</td>
<td>0.13</td>
<td>(0.77)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Selection procedure score</td>
<td>0.00</td>
<td>1.00</td>
<td>−0.08</td>
<td>−0.08</td>
<td>−0.27**</td>
<td>−0.20**</td>
<td>0.05</td>
<td>(0.85)</td>
<td></td>
</tr>
<tr>
<td>Situational interview</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Job relatedness</td>
<td>3.23</td>
<td>0.84</td>
<td>−0.07</td>
<td>−0.10</td>
<td>0.15*</td>
<td>0.50**</td>
<td>0.30**</td>
<td>0.00</td>
<td>(0.79)</td>
</tr>
<tr>
<td>8. Test-taking motivation</td>
<td>4.44</td>
<td>0.54</td>
<td>−0.03</td>
<td>0.02</td>
<td>0.20**</td>
<td>0.12</td>
<td>0.91**</td>
<td>0.04</td>
<td>0.30**</td>
</tr>
<tr>
<td>9. Selection procedure score</td>
<td>0.00</td>
<td>1.00</td>
<td>0.07</td>
<td>0.05</td>
<td>0.11</td>
<td>−0.16*</td>
<td>0.19**</td>
<td>0.34**</td>
<td>0.13</td>
</tr>
</tbody>
</table>

*Note.* N = 187. Selection procedure scores were standardized to obtain values with $M = 0.0$; $SD = 1.0$. Reliability estimates of the dependent variables are shown on the diagonal. Reliability estimates for job relatedness and test-taking motivation are coefficient alphas. Estimated reliability for the pencil-and-paper test selection procedure score is split-half reliability (Spearman-Brown corrected). Estimated reliability for the situational interview selection procedure score is an intraclass correlation. Reliability estimates were not calculated for the variables of candidate gender, organizational tenure, and race because they were single-item measures.

*p < 0.05; **p < 0.01.
comparative fit index (CFI) if 0.98, and a root mean squared error of approximation (RMSEA) of 0.06. All fit indices were well within acceptable ranges (Hu & Bentler, 1999).

Prior to testing the four hypotheses and assessing the two research questions, we conducted an omnibus test for the existence of overall differences in candidate reactions to and performance on the two types of selection devices. Therefore, a repeated measures multivariate analysis of variance (MANOVA) was conducted using perceived job-relatedness, test-taking motivation, and selection procedure score as the dependent variables. The “within” candidate factor was type of selection procedure (pencil-and-paper versus situational interview). The “between” candidate factor was promotion candidate race (African-American versus White). Although the study was primarily concerned with the association of these two factors with candidate perceptions (job-relatedness and test-taking motivation) and performance (selection procedure score), it is possible that factors other than race may be related to the three dependent variables. Thus, our analyses included employee gender and organizational tenure as control variables. These results revealed that gender and organizational tenure were not related ($p > 0.05$) to the dependent variables.

The repeated measures MANOVA identified differences for type of selection procedure, $F(3, 181) = 9.47, p < 0.001$, and promotion candidate race, $F(3, 181) = 6.44, p < 0.001$. The type of selection procedure $\times$ promotion candidate race interaction was also significant, $F(3, 181) = 11.05, p < 0.001$. The significance of the two main effects and interaction term in the repeated measures MANOVA indicated associations with at least one of the three dependent variables. Univariate repeated measures analyses of variance (ANOVA) were undertaken to test each of the individual hypotheses and examine each of the research questions.

**Type of selection procedure**

Hypothesis 1 posited that promotion candidates will perceive the situational interview as having greater job relatedness than the pencil-and-paper test. The repeated measures ANOVA showed that the promotion candidates perceived the situational interview ($M = 3.23, SD = 0.84$) as having higher perceived job-relatedness than the pencil-and-paper test ($M = 2.49, SD = 0.85$), $F(1, 183) = 24.46$, $p < 0.001$.
Thus, Hypothesis 1 was supported. As recommended by Hunter and Schmidt (1990), we used the pooled standard deviations in computing Cohen’s $d$ value. To adjust Cohen’s $d$ for the repeated measures nature of the study, we used the approach recommended by Nouri and Greenberg (1995) and Cortina and Nouri (1999). Effect sizes of 0.80 or greater are considered large effects, those around 0.50 are considered moderate, and those around 0.20 are considered small (Cohen, 1977).

Similarly, Hypothesis 2 stated that the promotion candidates will report higher test-taking motivation when completing the situational interview than when taking the pencil-and-paper test. The repeated measures ANOVA results showed that the type of selection procedure was not related to candidates’ test-taking motivation, $F(1, 183) = 0.01, p > 0.05$. Thus, contrary to our expectation, Hypothesis 2 was not supported. The promotion candidates viewed the two selection procedures similarly with respect to test-taking motivation.

**Promotion candidate race**

We also proposed that African-American promotion candidates would report higher perceived job-relatedness (Hypothesis 3a) and greater test-taking motivation (Hypothesis 3b) than White candidates. In terms of promotion candidate race, there were differences in candidate perceptions of job relatedness, $F(1, 183) = 7.62, p < 0.01, d = 0.44$ and test-taking motivation, $F(1, 183) = 11.95, p < 0.01, d = 0.47$. Specifically, African-American candidates gave higher job relatedness ratings of the selection procedures ($M = 3.00, SD = 0.92$) than did White candidates ($M = 2.72, SD = 0.93$). In addition, African Americans ($M = 4.55, SD = 0.52$) reported higher test taking motivation than Whites ($M = 4.31, SD = 0.56$). Therefore, both Hypotheses 3a and 3b were supported.

**Relationship of the interaction between type of selection procedure and promotion candidate race with candidates’ perceptions of the procedures**

We posed a research question: in an organization where the majority of leadership positions are occupied by African Americans, will promotion candidates’ race (African Americans versus Whites) interact with the type of selection procedures they complete (a situational interview versus a pencil-and-paper test) in reference to how they view these procedures (i.e., perceived job-relatedness and test-taking motivation)? No differences were found for the type of selection procedure × promotion candidate race interactions for either perceived job-relatedness or test-taking motivation.

**Relationship of the interaction between type of selection procedure and promotion candidate race with candidates’ performance on the procedures**

Our second research question was: in an organization where the majority of leadership positions are occupied by African Americans, will the interaction of promotion candidates’ race and type of selection procedure they complete be associated with candidates’ performance on the selection procedures?

Results of a repeated measures ANOVA indicated that the type of selection procedure × promotion candidate race interaction was significant for test performance, $F(1, 183) = 27.97, p < 0.001$. To illustrate the nature of the results, Figure 1 shows the mean standardized scores on each of the two selection procedures (pencil-and-paper and situational interview) separated by candidate race. Whites scored higher on the pencil-and-paper test than African Americans; however, African Americans scored higher than Whites on the situational interview. The standardized difference between the African-American and White scores on the pencil-and-paper test was $d = 0.56$. The standardized difference between the African-American and White scores on the situational interview was $d = 0.22$. 

Copyright © 2007 John Wiley & Sons, Ltd. 

J. Organiz. Behav. (in press) 
DOI: 10.1002/job
Discussion

The present study was conducted in an employment context in which African Americans represented the majority of the agency’s top-management members. We suggest that such an organizational context might be associated with racial differences in perceptions of pencil-and-paper tests and situational interviews used in making promotion decisions within such an organizational setting.

In regard to Hypothesis 1, we hypothesized that promotion candidates would perceive the situational interview as having greater job relatedness than the pencil-and-paper test. As predicted, promotion candidates viewed the situational interview as being more job-related than the pencil-and-paper job knowledge test. This result is consistent with previous research comparing perceptions of simulation exercises versus pencil-and-paper tests in entry-level selection contexts (cf., Macan et al., 1994; Schmidt et al., 1977; Smither & Pearlman, 1991; Smither et al., 1993).

Hypothesis 2 posited that promotion candidates would report higher test-taking motivation when completing a situational interview than when taking a pencil-and-paper test. Prior research has suggested that face validity perceptions affect test-taking motivation (Chan et al., 1997). Thus, we expected greater test-taking motivation for tests that were perceived as more job-related. However, this expectation was not confirmed; there was no difference between promotion candidates’ test-taking motivation for the two selection procedures. This result may have been because the context of this study was a highly competitive promotion process that represented a high-stakes testing condition. In support of this notion, it should be noted that test-taking motivation for both selection procedures was high ($M = 4.4$ on a 5-point rating scale) and range restricted (SDs for both procedures <0.60).

With regard to Hypotheses 3a and 3b, we hypothesized that African-American promotion candidates would (a) view the selection procedures as having higher job relatedness than would White candidates and (b) report higher test-taking motivation than would White candidates. As anticipated, we found that African-American and White candidates differed in their perceptions of job relatedness and test-taking motivation. African-American candidates had more positive perceptions of job relatedness and
test-taking motivation than White candidates did. Previous studies providing empirical evidence of racial differences in perceptions of job relatedness and test-taking motivation have produced mixed results. Schmit and Ryan (1997) and Smither et al. (1993), for example, reported no differences in test-taking motivation or face validity perceptions between African Americans and Whites. On the other hand, when racial differences in reactions have been identified, the findings typically have indicated lower levels of face validity and test-taking motivation for African Americans (Chan & Schmitt, 1997). In contrast, our finding that Whites provided lower ratings of job-relatedness and test-taking motivation than African Americans must be considered within the context of the high visibility of minority leadership in the organization.

Racial differences in real-world settings are highly influenced by context (Ryan & Ployhart, 2000). For example, African Americans have more positive views than Whites do about the fairness of testing in organizational settings where there are strong affirmative action programs and minorities in highly visible leadership positions (Ryan, Ployhart, Greguras, & Schmit, 1997; Schmit & Ryan, 1997). The present study involved promotion candidates employed in an urban police department that is racially diverse, with over half of the candidates being non-White. Additionally, the department's leadership was replete with minorities, especially in the upper-levels of management (e.g., the current and previous police chiefs are African Americans). These contextual factors may have played an important role in the nature of our findings (i.e., African Americans having more positive reactions than Whites did). Furthermore, our findings are noteworthy especially considering the range restriction normally associated with reactions-to-testing research. That is, only organizations with fair selection processes are inclined to allow a study of their selection procedures (Truxillo, Steiner, & Gilliland, 2004). As a result, in field studies of reactions to testing, mean ratings of perceptions of selection procedures are likely to be positively skewed with the most negative ratings being near the middle of a Likert rating scale (Truxillo, Steiner, & Gilliland, 2004). Therefore, the fact that this study found significant differences is important considering the ability to detect differences is limited by such range restriction.

With regard to our research questions, there was no evidence of a candidate race × selection procedure interaction in perceptions of job-relatedness or test-taking motivation. As discussed earlier, results of prior studies examining racial differences in perceptions of testing have been mixed (e.g., Chan & Schmitt, 1997; Chan et al., 1997; Chan, Schmitt, Sacco, et al., 1998; Rynes & Connerley, 1993). At least for a pencil-and-paper test and situational interview, our findings showed that promotion candidates’ race did not interact in such a way that African-American and White candidates’ perceptions of job-relatedness and test-taking motivation varied depending on the selection measure employed. As we have proposed, organizational context may account for these results. However, the results of this study must be interpreted with caution as it is based on a study of a single organization; further research in organizations (especially those with varying levels of African-American top leadership) is needed.

The finding that the test-taking motivation of African Americans and Whites did not differ as a function of exercise type is of particular interest. Many researchers have suggested that differences in test-taking motivation may be the reason for performance differences between African Americans and Whites on standardized tests, and that this performance difference can be reduced if test-taking motivation of African Americans is improved (McKay & Doverspike, 2001). One of the more popular approaches for improving test-taking motivation is to use different testing formats that are thought to stress the values and culture of the African-American culture (i.e., oral and performance tests). Our results suggest that such an approach may be ineffective for this purpose in the present organizational context, as there was no difference in the test-taking motivation regardless of whether the exercise was verbal (i.e., situational interview) or written (i.e., pencil-and-paper test) in nature.

In regard to our second research question (i.e., the relationship of candidate race × type of selection procedure with selection procedure performance), African Americans outperformed White candidates on the situational interview while White candidates outperformed African-American candidates on the
pencil-and-paper test. This is an interesting result since African-American candidates reported higher levels of test-taking motivation than White candidates for both situational interviews and pencil-and-paper tests, yet African-American candidates performed more poorly than Whites on the paper-and-pencil test. This finding would seem to indicate that, contrary to many researchers’ suggestions, lower test-taking motivation may not necessarily be the reason for racial differences in performance on pencil-and-paper tests. The performance gap between African Americans and Whites has been one of the most vigorously researched topics in education, industrial psychology, and human resource management and continues to be one of the most pressing social issues. The African-American—White test performance gap has also received much public attention, especially after the publication of works such as *The Bell Curve* (Herrnstein & Murray, 1994) and stereotype threat theory (Steele & Aronson, 1995). In essence, the African-American community is well aware of this problem due to the publicity it has received, and this awareness affects the way African Americans perceive and possibly perform on such tests (Sackett, Hardison, & Cullen, 2004). Therefore, many argue that this knowledge of underperformance as a group on such tests is the primary reason for the African-American—White performance difference (cf., Steele & Aronson, 1995). In our study, African-American candidates viewed both structured interviews and pencil-and-paper tests favorably, even more favorably than White candidates. Yet, African-American candidates still underperformed when compared to White candidates on the pencil-and-paper test despite the fact that they viewed them more favorably. This would seem to provide evidence against such an explanation of the test performance gap. Other recent studies, Cullen, Hardison, and Sackett (2004), for example, have produced results challenging this contention as well. Our results suggest that, while a more diverse organizational leadership may improve perceptions of job-relatedness and test-taking motivation among African Americans, such an organizational composition may not alleviate racial performance differences on pencil-and-paper tests.

**Study limitations and directions for future research**

One limitation of our study is that test method and test content were confounded. Without controlling for test content across test format (cf. Chan & Schmitt, 1997), the effects of test format cannot be separated from the effects of test content and *vice versa*. Chan and Schmitt (1997) suggest that one way of separating the two different effects is to examine a common set of test items across different methods of testing. Chan and Schmitt also state that this may be extremely difficult for some constructs and methods. One way to achieve this in future research would be to administer a pencil-and-paper test and a situational interview that measure the same constructs or KSAs. Because this study was conducted during a *bona fide* promotion procedure in an actual organization, this approach would have been prohibitively expensive in terms of development and administration costs and could have exposed the organization to unnecessary legal risks. Since these were significant costs and risks to the organization, such an approach was not feasible or appropriate. Therefore, while this limitation is recognized, it is often a necessary tradeoff in conducting research in actual selection situations using *bona fide* job candidates, especially since field research is the most meaningful strategy for improving our understanding of reactions to testing (Smither et al., 1993).

Another potential study limitation concerns effect size. Whether using $\eta^2$ or $d$ as an index of effect size, one could argue that because of the small to moderate effect sizes, these differences may have no real practical significance. However, it is equally important to recognize that African Americans did not react more negatively than Whites to the exercises, especially the pencil-and-paper test. Even if the current study had found no significant differences in the reactions of African Americans and Whites, it would have been practically important because the result would have been directly opposite from previous findings (*i.e.*,...
African Americans having more negative perceptions of pencil-and-paper tests), and we suggest this result might be due to the organizational context. We theorize that the racial composition of top management in the organization may have had an effect on African-American candidates’ reactions to the tests. We reason that this result may be due to the similar-to-me effect or social identification theory. Because of the large percentage of African Americans in the organization’s top management, African-American candidates may have seen top management as similar to themselves and trustworthy. As a result, African-American candidates may have judged the selection process, including the pencil-and-paper test, more favorably than in other studies due to their positive perception of top management.

In addition, another limitation of our study concerns statistical power. Statistical methods such as moderated multiple regression and repeated measures MANOVA commonly used in testing hypotheses involving moderator variables can have limited power to detect interaction effects (Aguinis, 1995). Therefore, it is possible that candidate race × test type interactions may be associated with candidate selection procedure reactions, but we simply lacked the power to detect these relationships. Future research should incorporate larger samples in order to provide a more adequate test of candidate race × test type interactions (Aguinis, 1995).

Finally, our study was not multi-organizational, but rather involved one organization with a strong presence of minority leadership throughout the organization. Research in different organizational contexts is needed before drawing any firm conclusions about the influence of particular aspects of organizational context. Therefore, an interesting area for future research is examining reactions to assessment in different organizational contexts (e.g., ones having different percentages of African Americans in leadership positions).

Author biographies

**J. Bret Becton** is an Assistant Professor of Management in the Management and Marketing Department at Winthrop University. He earned his Ph.D. in Human Resource Management from Auburn University. His research interests include group differences in performance on selection procedures, candidate reactions to selection procedures, organizational citizenship behavior, executive compensation, and organizational justice. His research has appeared in *Public Personnel Management, Coastal Business Journal, and Public Management* and he has consulted with numerous organizations in both the public and private sector.

**Hubert S. Feild** is Torchmark Professor of Management in the College of Business at Auburn University. He earned his Ph.D. in Industrial/Organizational Psychology from the University of Georgia. His research interests include human resource selection and research methods in human resource management. He has published articles in numerous journals such as the *Journal of Applied Psychology, Personnel Psychology*, and the *Academy of Management Journal* and has consulted with organizations such as SONY, GE, and PPG Industries. His most recent book is *Human Resource Selection* (Harcourt) that he coauthored with Robert Gatewood.

**William F. Giles** is a Professor of Management at Auburn University. He received his Ph.D. in Industrial/Organizational Psychology from the University of Tennessee. His primary research interests are the area of performance appraisal, with secondary interests in training and selection. His research has been published in journals such as *Journal of Applied Psychology, Academy of Management Journal*, and *Personnel Psychology.*

**L. Allison Jones-Farmer** is an Associate Professor of Management at Auburn University. She received her Ph.D. in Applied Statistics from the University of Alabama. Professor Jones-Farmer’s research
interests include Statistical Process Control, Control Charts, and Multivariate Methods. Her research has been published in several scholarly journals including *Technometrics*, *Journal of Quality Technology; Quality and Reliability Engineering International* and *Journal of Statistical Computation and Simulation*.

**References**


Copyright © 2007 John Wiley & Sons, Ltd.

J. Organiz. Behav. (in press)

DOI: 10.1002/job


