

9-1-2015

## **Interviewee Selection Test and Evaluator Assessments of General Mental Ability, Emotional Intelligence and Extraversion: Relationships with Structured Behavioral and Situational Interview Performance**

Donald Kluemper

Benjamin McLarty

Terrence R. Bishop

Anindita Sen

Follow this and additional works at: <https://huskiecommons.lib.niu.edu/allfaculty-peerpub>

---

### **Original Citation**

Kluemper, Donald, et al. (2015). Interviewee Selection Test and Evaluator Assessments of General Mental Ability, Emotional Intelligence and Extraversion: Relationships with Structured Behavioral and Situational Interview Performance. Vol. 30 (Issue 3) pp. 543-563. DOI: 10.1007/s10869-014-9381-6.

This Article is brought to you for free and open access by the Faculty Research, Artistry, & Scholarship at Huskie Commons. It has been accepted for inclusion in Faculty Peer-Reviewed Publications by an authorized administrator of Huskie Commons. For more information, please contact [jschumacher@niu.edu](mailto:jschumacher@niu.edu).

# Interviewee Selection Test and Evaluator Assessments of General Mental Ability, Emotional Intelligence and Extraversion: Relationships with Structured Behavioral and Situational Interview Performance

Donald H. Kluemper · Benjamin D. McLarty ·  
Terrence R. Bishop · Anindita Sen

Published online: 26 September 2014  
© Springer Science+Business Media New York 2014

**Abstract** The purpose of the study was to examine antecedents of interview performance commonly measured via two divergent methods; selection tests and evaluator assessments. General mental ability (GMA), emotional intelligence (EI), and extraversion have been largely studied in isolation. This study evaluates the relative strength of these traits across methods and tests whether selection test and evaluator-assessed traits interact to further enhance the prediction of interview performance. 81 interviewees were asked to complete traditional selection tests of GMA, EI, extraversion, and a video-recorded structured behavioral and situational job interview. The traits, behavioral, and situational interview performance were then evaluated with three independent sets of raters. Regression analysis was used to investigate the extent that these traits predicted structured interview performance. Results indicate that each trait was a strong predictor of interview performance, but results differed based on the method of measurement and the type of structured interview assessed. Further, evaluator perceptions related to interview performance more strongly than did selection

tests. Finally, evaluator assessments of each trait interacted with its respective selection test counterpart to further enhance the prediction of interview performance. This improves our understanding of how applicant traits impact hiring decisions. This is the first study to directly compare tested versus others' ratings of interviewee GMA, EI, and extraversion as predictors of interview performance.

**Keywords** Behavioral descriptive interview · Extraversion · Emotional intelligence · General mental ability · Personality

## Introduction

Despite nearly a century of published research on employment interviewing, little is known about the factors that affect how an interviewee performs in an employment interview (Huffcutt et al. 2011). As presented by Huffcutt et al., there are multiple reasons for this lack of knowledge. Past research has focused on aspects of interview ratings such as criterion-related validity (Huffcutt and Arthur 1994), reliability (Conway et al. 1995), subgroup differences (Huffcutt and Roth 1998), and incremental prediction beyond other predictors of job performance such as general mental ability (GMA; Cortina et al. 2000).

In an attempt to address this issue, Huffcutt et al. (2011) present a new theoretical perspective that distinguishes interview performance from interview ratings, arguing that studying interview performance more directly is vital to advance this stream of research. They define interview performance as how applicants behave during the interview, what they say and what they do, including the content of interview verbalizations, how they deliver content and nonverbal behavior. By advancing research using this

---

Portions of this manuscript were presented at the Southern Management Association annual conference in 2009 in Asheville, NC.

---

D. H. Kluemper (✉)  
Department of Managerial Studies, University of Illinois at  
Chicago, 601 S. Morgan St., 2204 UH, Chicago, IL 60607-7122,  
USA  
e-mail: kluemper@uic.edu

B. D. McLarty · A. Sen  
Louisiana State University, Baton Rouge, LA, USA

T. R. Bishop  
Northern Illinois University, DeKalb, IL, USA

perspective of interview performance, a renewed theoretical attention on the myriad of individual differences involved can be examined. For example, general mental ability (GMA; Hunter and Hunter 1984), emotional intelligence (EI; O’Boyle et al. 2010), and personality traits such as extraversion (Hurtz and Donovan 2000) can all be examined to see the ways in which interviewees present their qualifications during an interview process.

Prior studies investigating the influence of individual differences on interview performance have utilized different approaches to measure these predictors. These studies have used either traditional trait selection methods administered to the interviewee, such as tests of GMA (Salgado and Moscoso 2002) and self-assessed personality tests (Huffcutt et al. 2001a), hereafter referred to as “selection tests.” Other interview studies have used evaluator ratings of the interview to assess traits such as EI (Sue-Chan and Latham 2005) and personality (Roth et al. 2005). In these studies, trained raters observe the interview for the purpose of rating particular interviewee traits exhibited during the interview, which we hereafter refer to as “evaluator assessments.”

Our study seeks to contribute to ongoing selection research by simultaneously examining GMA, EI, and extraversion as assessed through both selection tests and evaluator assessments, as well as the potential interactions between selection tests and evaluator assessments. These issues will now be addressed in greater detail.

### Selection Tests Versus Evaluator Assessments

Evaluator assessments may tap different aspects of traits than that which is measured through selection tests. Socioanalytic theory (Hogan 1996, 2007) proposes that individual differences can be viewed from two separate but related perspectives. Hogan labels these terms *identity* and *reputation*. Identity is an individual’s self-assessment, meaning what one internally believes about themselves and their idealized vision of who they are. Identity influences agendas that individuals engage in and what ways they will play various roles (Hogan and Shelton 1998). Self-assessments are one of the most common forms of assessment used to determine individual differences (Oh et al. 2011) such as one’s identity.

On the other hand, identities are translated into one’s reputation according to the self-presentational style that different individuals enact (Hogan 2007). Thus, reputation is described as one’s interpersonal style and the observable actions associated with it (i.e., the impression one gives off). Prior researchers have reliably assessed individual differences from an observer perspective (e.g., Funder and Sneed 1993; Kolar et al. 1996) demonstrating the

observer’s evaluation of a focal actor’s behaviors, expressed beliefs, desires, and motives. Therefore, reputation appears to be stable over time and highlights the importance of using past performance to predict future behavior—including interview performance (Hogan 2007; Mount et al. 1994). In combination, these differing perspectives suggest that individual differences should be seen from the perspective of both an observer and the actor (Hogan 1996).

In this vein, selection tests of personality incorporate less observable information about motives, intentions, and feelings, whereas evaluator ratings are tied to observations of target behaviors. Thus, from a socioanalytic theory perspective, internal aspects of personality (i.e., identity) are distinct from evaluator-assessed aspects of personality (i.e., reputation). Therefore, this framework parallels the measurement differences that are highlighted using selection tools and evaluator assessment methods. Reputational ratings (i.e., evaluator methods) may indeed represent “reality” as past performance (i.e., reputation) can accurately predict future performance (Hogan 2007; Mount et al. 1994).

Recent work has established stronger criterion-related validity coefficients for other assessments over self-assessments for each of the Big Five personality traits (Oh et al. 2011) and for EI (Choi and Kluemper 2011). Further, other ratings are particularly valuable when self-assessments are untrustworthy and when researchers wish to improve accuracy by aggregating multiple raters (Hofstee 1994; McCrae and Weiss 2007). Although a majority of studies investigating self-assessments versus other assessments have focused on assessments from acquaintances with a certain degree of familiarity of the individual being assessed, recent research has begun to establish a variety of contexts in which unacquainted evaluators can accurately assess the traits of those they observe. Examples include personality assessment through the evaluation of word use (Fast and Funder 2008), email (Gill et al. 2006), resumes (Cole et al. 2003, 2005), attire (Burroughs et al. 1991), and even social networking profiles like Facebook (Kluemper and Rosen 2009; Kluemper et al. 2012). Further, self- and peer-assessed personality rely on memory recall, which introduce various biases (Highhouse and Bottrill 1995; Srull and Wyer 1989) not present when traits are obtained through evaluator assessments. Thus, grounded in socioanalytic theory, we suggest that trait visibility in the interview may be of sufficient quantity and quality as to influence interview performance.

Huffcutt et al. (2011) argue that highly structured interviews are the most appropriate context in which to assess a more complete and accurate effect of interviewee traits (for an opposing view, see Blackman 2002). First, structured interviews developed from job analysis focus the

content of the interview on core candidate qualifications, priming interviewees to more readily present job-relevant traits. Second, structure serves to reduce the influence of interviewer–interviewee dynamics, thereby reducing rater bias and error. Thus, more structured interviews should more closely align interview performance both with job-relevant interviewee qualifications and interview ratings.

Along these lines, we focus on highly structured interviews as the most appropriate context in which to assess the impact of interviewee traits on interview performance. In fact, Van Iddekinge et al. (2005) found that personality (altruism, self-discipline, and vulnerability) could be reliably assessed by others in a personality-based employment interview. Thus, traits assessed via traditional assessment methods and perceptions of traits on the part of an evaluator should relate to interview performance.

We focus on structured behavioral interviews to assess GMA, EI, and extraversion (Janz 1982) because they are “more verbally intensive” than other types of structured interviews (Huffcutt et al. 2001b, p. 624). In a structured behavioral interview, there are two important sources of trait visibility. First, the interview setting itself provides a source, through observable behaviors typically associated with a wide range of interpersonal interactions. As such, the interviewer observes nonverbal cues such as eye contact, body orientation, smiling, gesturing, (DeGroot and Gooty 2009), vocal characteristics (DeGroot and Motowidlo 1999), and word use (Fast and Funder 2008). Second, the behavioral interview utilizes past-oriented questions by describing a situation and asking respondents how they have behaved in the past in such a situation (McDaniel et al. 1994), yielding more behavioral content (Janz) and incidental measurement of general characteristics (Huffcutt et al. 2001a). The nature of behavioral interviews requires candidates to provide information on the context and dynamics of each past experience before providing detailed information regarding their behavior in the situation being recalled (Huffcutt et al. 2001b). The interviewees’ personality-relevant behaviors that occurred during a past incident are recalled and reported to the interviewer, which adds to the trait visibility of behavioral interviews. Thus, trait perceptions on the part of the evaluator will result from legitimate situation-specific behaviors and recall of personality-relevant past behaviors, which will influence interview performance ratings.

## Hypothesis Development

### General Mental Ability

Since the very earliest research on personnel selection, GMA has been one of the major methods used to

distinguish between candidates and to predict subsequent job performance across more or less all occupational areas (Robertson and Smith 2001). GMA is the single most effective predictor of job performance, accounting for approximately 15 % of variance for unskilled jobs to approximately 53 % of variance for complex jobs, when correcting for range restriction (Hunter et al. 2006). Although valid tests of GMA exist, GMA is often assessed in interviews for a variety of reasons, including logistic considerations, habit, and legal considerations (Huffcutt et al. 2001a). Huffcutt et al. (1996) propose that highly structured interviews could be measuring cognitive factors such as GMA. In fact, Hunter and Hirsch (1987) postulate that structured interviews operate as oral tests of GMA. Several meta-analyses demonstrate a moderate to high corrected correlation between GMA and interview ratings. Specifically, Roth and Huffcutt (2013) found a correlation of .42, Huffcutt et al. (1996) found a correlation of .40, and Berry et al. (2007) found a correlation of .27.

In line with the arguments above regarding observer ratings, GMA should foster the trait visibility necessary to affect interview performance. To support this claim, evidence exists for the accurate assessment of GMA by observers (Borkenau et al. 2004). The interview provides a broader variety of cues and data from which an evaluator might make judgments. This information includes the interviewees’ vocabulary, gestures, grammatical skill, information depth, and other data.

Further, behavioral interviews require the interviewee to describe the context and dynamics of each past experience, how they responded, and to summarize the outcomes of their actions. Thus, the cognitively demanding nature of behavioral interviews should result in enhanced trait visibility of GMA, such as a more comprehensive recall of the context, experience, and outcomes of the past event. In conjunction with issues related to the cognitively demanding nature of behavioral interviews, some interviewees have the capability (i.e., those high in GMA) to understand the underlying nature and intent of interview questions (Konig et al. 2007). In addition, as GMA is a strong predictor of job performance (Hunter and Hunter 1984), the behavior of the interviewee in the past situation being recalled should indicate superior performance in that past situation when that interviewee is high in GMA.

Perceptions of interviewee intelligence on the part of the interviewer should, in turn, equate to elevated interviewer judgments of interview performance. In support of this argument, Dunn et al. (1995) conducted a policy-capturing study indicating that hiring managers generally view GMA and personality traits as important for the hirability of the applicant, even though these hiring managers were typically unaware that they were utilizing this information. Further, hiring managers are influenced more when they

know this information is obtained via interviews rather than from employment selection tests (Lievens et al. 2005). In addition, GMA has been found to relate to interview performance ratings with tests of GMA (Campion et al. 1988) and through interviewer assessments (Huffcutt et al. 2001a), although results connecting GMA to interview ratings are mixed (Robertson and Smith 2001; Salgado and Moscoso 2002). Thus

**Hypothesis 1** General mental ability measured via selection test (1a) and evaluator assessment (1b) will be positively related to interview performance.

#### Emotional Intelligence

EI is argued to have important implications for the selection of employees in organizations (Fisher and Ashkanasy 2000; Choi et al. 2011; Kluemper et al. 2013). Given the interpersonal (Huffcutt et al. 2011) and emotionally charged nature (Ashkanasy et al. 2002) of the employment interview, it is likely that EI would predict interview performance. Interviewees are energized and aroused to present the image of a potentially effective employee. Conversely, they may be adversely affected by anxiety, such as the fear they might not be able to answer the interview questions “correctly,” or the fear that they will otherwise be seen to behave inappropriately in the interview. Individuals with EI are more socially effective than others (Salovey and Mayer 1990) and should get interviewers to like and feel good about them by effectively utilizing empathy, self-presentation, and tactical use of nonverbal expression (Fox and Spector 2000).

Further, individuals high in EI tend to be more effective at recognizing and regulating their own moods and feelings. In a job interview, a candidate in a positive mood may be more likely to recall, construe, and describe incidents of past work performance in a self-enhancing way, may be more likely to project a confident and competent self, and may be more adept at dealing creatively with unexpected questions (Fox and Spector 2000). In this vein, Huffcutt et al. (2001a) postulate that behavioral ratings are influenced by the social skills of the interviewee, while Baron’s (1993) work has shown that emotional competence is important in the interview. Further, interviewees who express positive affect and are empathetic are likely to be more successful in generating positive impressions in the interviewer (Fox and Spector 2000). As knowledge beyond these findings is limited, it would prove beneficial to further investigate the effect of the interviewee’s specific emotional intelligence on interview performance.

Thus, EI-related trait visibility is likely present in the behavioral interview, again, creating the conditions necessary for the assessment of reputation-related personality

according to socioanalytic theory. In fact, recent work by Sue-Chan and Latham (2005) found that evaluator-assessed EI correlated at .31 to highly structured situational interview ratings. Due to obvious similarities between structured situational interviews and the behavioral descriptive interview, it stands to reason that emotional intelligence assessed in behavioral interviews will predict interview performance as well, particularly given the more verbally intense nature of the behavioral descriptive interview when compared to the situational interview (Huffcutt et al. 2001b).

**Hypothesis 2** Emotional intelligence measured via selection test (2a) and evaluator assessment (2b) will be positively related to interview performance.

#### Extraversion

Among the popular Big Five personality framework (Digman 1990), extraversion is particularly relevant in job interviews. Extraverted individuals exhibit more energy and enthusiasm than their introverted counterparts (John and Srivastava 1999), expressions that will likely convey interest in the position and a higher level of motivation on the job, and thus will result in a more favorable impression by the interviewer. Individuals high in extraversion also exemplify characteristics such as sociability and talkativeness (McCrae and Costa 1999), resulting in longer and more detailed answers to behavioral questions. Further, extraverts are more assertive than are introverts (McCrae and Costa 1999), a trait that will likely foster favorable interviewer perceptions of esteem and efficacy. These aspects of extraversion are also likely to be exhibited in the past behaviors recalled and conveyed via behavioral interviews.

Given these characteristics, as informed through the tenets of socioanalytic theory, we can expect that extraversion will enhance interview performance. In fact, Kristof-Brown et al. (2002) found that extraverted interviewees had a greater tendency to use self-promotion. In other work, Caldwell and Burger (1998) noted that extraversion is probably the most important personality trait during the interview interaction and found that it influenced interview decisions in their study. Likewise, Salgado and Moscoso (2002) found a meta-analytic link between behavioral interview ratings and personality test scores for extraversion (.21), which was the strongest predictor of interview ratings among the Big Five personality traits. Thus, we propose the following hypothesis:

**Hypothesis 3** Extraversion measured via selection test (3a) and evaluator assessment (3b) will be positively related to interview performance.

## Relative Strength and Interactive Effects

Our theorizing and prior research findings support relationships between GMA, EI, and extraversion, and interview performance. Less clear, however, is which of these three traits, when assessed simultaneously, is the strongest predictor. Further, due to differences discussed earlier between selection tests and evaluator ratings, the relative strength of these traits may differ across these divergent measurement methods. Thus, the simultaneous assessment of both selection tests and evaluator assessments of all three individual differences (GMA, EI, and extraversion) will allow a more comprehensive understanding of their relative strength in predicting interview performance. Based on the theoretical rationale provided above for GMA, EI, and extraversion, along with the established magnitude of the relationships found in prior research, we posit that GMA will be the strongest predictor of interview performance, followed by extraversion, and finally EI.

**Hypothesis 4a** Among selection test measures, GMA will be the strongest predictor of interview performance.

**Hypothesis 4b** Among evaluator assessment measures, GMA will be the strongest predictor of interview performance.

In line with socioanalytic theory, evaluator assessments (i.e., reputation) should have a stronger impact on interview performance than selection tests (i.e., identity). Because evaluator assessments of individual differences are a measure of past performance and past performance is strongly predictive of future performance (Hogan 2007; Mount et al. 1994), it follows that this form of assessment should relate more strongly to consequent interview performance. Thus, the three evaluator-assessed traits are likely larger than their respective selection test counterpart in the prediction of interview performance.

**Hypothesis 4c** The strength of the evaluator-assessed traits will be higher than the strength of corresponding traits measured via selection tests.

Traditional self-assessed traits (i.e., extraversion and EI) represent a respondent's thoughts, such as schemas, norms, and expectations, feelings such as affect and arousal, and recall of one's general behavioral tendencies (Roberts 2009). Similarly, GMA represents a person's overall mental capacity through effective cognition and information processing (Brody 1992). Inherent in these conceptualizations is that certain behaviors are, in general, more likely to result from individuals with higher levels of these traits. As explained by socioanalytic theory (Hogan 1996, 2007), these individual differences are expressions of an individual's identity (extraversion and EI) and ability

(GMA). On the other hand, evaluator assessments represent the reputation that an individual has developed regarding their specific observable behaviors. Therefore, socioanalytic theory explicates how evaluator-rated assessments of individual differences are unique from self-assessments, and how these two forms could be combined into an interactive model for enhanced interview performance assessment. Whereas identity explains *why* someone behaves a certain way, reputation details *what* an individual has done (Hogan 2007). This dual interpretation of individual difference assessment gives differing perspectives regarding the validity of predicting work-relevant outcomes such as interview performance. Self-assessments and evaluator assessments do not capture identical constructs; however, they are meaningfully associated and valuable for predicting future behavior (Oh et al. 2011).

Because socioanalytic theory argues for the existence of both identity and reputation as different aspects of personality, the different methods used to assess personality can provide us with potentially complimentary information. For example, when reputation ratings (i.e., evaluator ratings) and identity ratings (i.e., self-ratings or selection tests) are both low for a specific individual difference, then it is very likely that the specific trait in question is actually low for that person (e.g., low extraversion). Therefore, if the individual difference (GMA, EI, or extraversion) is largely absent, both in one's identity (i.e., selection test) and reputation (i.e., evaluator assessment), then the person in question should not perform well in the interview. Further, recent research shows self- and other-rated traits such as conscientiousness and agreeableness can interact with one another to predict workplace outcomes (Kluemper et al. 2014). As such, we posit that the lowest levels of interview performance will result when both selection tests and evaluator assessments are low. Therefore, we hypothesize the following:

**Hypothesis 5** Selection tests of general mental ability (5a), emotional intelligence (5b), and extraversion (5c) will interact with their respective evaluator assessments to predict interview performance, such that low levels of both selection test and evaluator assessments will yield the lowest levels of interview performance.

## Method

### Participants and Procedures

Our sample consisted of 81 participants currently employed in the position of Youth Treatment Specialist at a residential treatment center in the Midwestern U.S. Youth Treatment Specialists are responsible for direct supervision

of approximately 12 delinquent juveniles. In terms of demographic characteristics, 51 % of participants were female, 73 % Caucasian, and were 25.0 years of age on average (ranging from 19 to 51 years). Participants had an average of 4.8 years of work experience and consisted of 13 % with only a high school diploma or GED, 63 % with some college, and 25 % with at least a bachelor's degree.

Participants volunteered for in-person sessions to complete a battery of selection tests (GMA, EI, and extraversion) and a video-recorded structured interview. The video-recorded interview consisted of a series of behavioral and situational questions. Upon the conclusion of the interviews, these behavioral and situational interview segments were then separated into independent video clips for evaluation. Three independent sets of three evaluators assessed these clips. One set of evaluators assessed the traits from interviewee responses to the behavioral interview questions, and a different independent set of evaluators assessed interview performance on this same set of behavioral interview questions. Accordingly, both independent sets of evaluators based their evaluations (traits; performance) on the same source (interviewees' responses to the behavioral interview questions). To mitigate same source bias, we had a third independent set of three evaluators provide interview performance ratings only for the situational interview question set embedded within the interview. The interview consisted of 10 questions, 5 situational, and 5 behavioral.

#### *Structured Interview Development and Administration*

Prior to the development of the behavioral interview in this study, a job analysis was conducted for the position of "Youth Treatment Specialist." The primary researcher (a former incumbent and supervisor) developed a thorough list of 18 task statements. These task statements were derived from internal (organizational specific documentation such as the job description, performance evaluation, and training documentation) and external (O-NET and related external job descriptions) analysis. This list of task statements was then evaluated and rank ordered by 20 incumbents and supervisors of the organization. The performance criteria found to be consistently identified as important for job success (leadership, initiative, persuasiveness, thoroughness, and oral communication skill) were then formed into questions (Janz 1982; Latham et al. 1980), thus enhancing content, interrater, and test-retest reliability (Campion et al. 1997). Five behavioral descriptive and five situational questions (included in the "Appendix" section) were developed to assess these performance criteria. Behavioral and situational question were administered in alternating order, such that half of the interviewees had the five behavioral questions first, while the other half received

the five situational questions first. The video-recorded interviews were then split into two independent video clips so that they could be assessed by different evaluators. This approach allows us to minimize the concern that the trait and performance ratings may be inflated due to both sets of raters evaluating the same behaviors. Interviews averaged 18 minutes in length.

The questions used in this study contain well-developed and detailed anchored scales, illustrating a 7 (high), a 4 (moderate), and a 1 (low); along with multiple behavioral anchors for each question. This highly structured multiple anchor approach is expected to increase validity, test-retest reliability, and interrater reliability (Campion et al. 1997). Using the same interviewer is very important in increasing structure because different interviewers may ask different questions and ask the questions differently (Campion et al. 1997). Variance due to interactions with candidates should be reduced due to less variation in interviews. Therefore, the same interviewer (the first author) conducted all job interviews. Frequent prompts and follow-up questions are a primary means by which interviewers might bias information gathering (Dipboye 1994). Structured interviews that minimize or omit probing show more robust and consistent reliability and predictive validities than do structured interviews that make more frequent use of probing (Huffcutt and Arthur 1994; Campion et al. 1997). Therefore, prompting, follow-up questions, and elaboration on questions were minimized. Prompting was used when the interviewee's response was too brief, when the interviewee was not answering the question given, or when the interviewee needed clarification on a question.

Uncontrolled questions from candidates reduce standardization by changing the interview content in unpredictable ways. Not allowing questions from candidates should standardize the content, thus increasing test-retest and interrater reliability (Campion et al. 1997). It prevents interviewers from using candidate questions to judge candidates, and it prevents candidates asking questions and using the information to shape their answers (Beatty 1986). Therefore, as is commonly done in structured interview formats, interviewees had an opportunity to ask questions at the end of the interview, allowing the opportunity to omit that segment from the video recordings given to the raters. A threat to structure is the uncontrolled use of ancillary information including application forms, resumes, test scores, recommendations, previous interviews, transcripts, and so forth. It confounds the interpretation of the value of the interview. Withholding this information should increase test-retest and interrater reliability (Campion et al. 1997) and validity. Therefore, neither the interviewer nor the evaluators had access to ancillary information including resumes, the GMA test, personality assessment, the EI test scores, etc.

### Evaluator Assessment of Traits

After filming and editing was complete, a group of three evaluators then viewed each of the 81 video-recorded behavioral descriptive interviews to assess GMA, EI, and extraversion. The evaluators were graduate assistants from a large Southern university. After a two hour training session, the evaluators met twice per week for approximately two hours per session until the trait assessments were complete. The training session consisted of a review of the extraversion, GMA, and EI constructs and definitions, appropriate utilization of rating scales, familiarization with the rating forms to be used, and practice conducting trait ratings for two pilot video-recorded behavioral descriptive interviews.

Multiple raters were utilized because they may be beneficial for several reasons. Multiple raters may reduce the effect of idiosyncratic biases among raters (Campion et al. 1988; Hakel 1982), and aggregating multiple judgments cancels out random errors (Dipboye 1992; Hakel 1982). The range of information and judgments from different perspectives may increase convergent validity (Dipboye 1992). Finally, using more raters is akin to a longer test; thus, the combined scores should be more reliable (Hakel 1982). Internal consistency should be higher because more judgments make up the total scores (Campion et al. 1997).

### Interview Performance Ratings

Behavioral interview performance was assessed by three graduate assistants from a large university in the Southwest U.S. In addition, situational interview performance was assessed by three different graduate assistants from the same university. These evaluators were not the same individuals who evaluated the traits described above. Using independent evaluators for IVs and DVs allows for the reduction of common-method variance (Podsakoff et al. 2003) across traits and behavioral and situational interview performance ratings. These raters also participated in a two hour training session. The training session consisted of familiarization of the job description in which the interviews were based, familiarization with the structured interview questions, emphasis on notetaking, a review of structured interview assessment, appropriate utilization of rating scales, familiarization with the structured rating scales used in the study, and a practice session of structured ratings for two pilot videos of recorded interviews.

Evaluators met once per week for approximately two hours per session until the interview ratings were complete. The evaluators watched the video-recorded interviews while taking notes. Notetaking may enhance structure because it reduces memory decay (Campion et al. 1988) and avoids recency and primacy effects (Schmidt and Ostroff 1986). Notetaking should also make evaluations more

consistent, thus increasing validity, test–retest, and inter-rater reliability (Burnett et al. 1998). Discussing candidates may lead to irrelevant information entering the evaluation process, thereby decreasing the validity of the interview (Campion et al. 1997). Therefore, raters did not communicate with one another throughout the rating process. Immediately following the viewing of each interview, the evaluators completed the structured interview ratings of the five performance dimensions (leadership, initiative, persuasiveness, thoroughness, and oral communication skill). Conducting ratings at the end of the interview is less structured than conducting one rating after each question, but more structured than conducting an overall rating at the end of the interview (Campion et al. 1997). Each of the rater scores was averaged to form a composite for each interview dimension. These rating dimensions were then averaged to produce interview performance ratings for the behavioral and situational interviews.

### Measures

To measure selection test extraversion, we used the 12-item NEO-Five-Factor Inventory (Costa and McCrae 1985). A sample item for extraversion is “I am a cheerful, high spirited person” rated on a 5-point scale ranging from *strongly disagree* to *strongly agree*.

To assess EI, we used the Wong & Law Emotional Intelligence Scale (WLEIS). Wong and Law’s (2002) EI measure is 16-item scale based in Mayer and Salovey’s (1997) model. Respondents were asked to complete questions such as “I have good understanding of my own emotions” on a 7-point scale ranging from *strongly disagree* to *strongly agree*.

We measured GMA with the Wonderlic Personnel Test, which has 50 questions with alternate form reliability coefficients that range from .73 to .95 and test–retest reliability coefficients ranging from .82 to .94 (Wonderlic and Associates 1992).

Impression management was included as a control variable and measured with 6 items from Reynolds (1982). A sample item is “I’m always courteous, even to people who are disagreeable” rated on a 5-point scale ranging from *strongly disagree* to *strongly agree*.

To measure evaluator ratings of extraversion, we used the Ten-Item Personality Inventory (Gosling et al. 2003), consisting of 2 extraversion items. Items for extraversion include “reclusive/sociable” and “reserved/quiet.”

We measured evaluator ratings of EI with two items based on the Wong and Law’s (2002) scale. These items were modified to be assessed by evaluators and include “Able to perceive emotions in self and others” and “Able to adequately express emotions.”

**Table 1** Descriptive statistics and intercorrelations for study variables

Variable	1	2	3	4	5	6	7	8	9
1 ST Extraversion									
2 ST Emotional Intelligence	.40*								
3 ST General Mental Ability	.15	.12							
4 EA Extraversion	<u>.36*</u>	.18	.24*						
5 EA Emotional Intelligence	<u>.28*</u>	<u>.22*</u>	.09	.59*	.43*	.36*			
6 EA General Mental Ability	<u>.21<sup>^</sup></u>	<u>.04</u>	<u>.27*</u>	.58*	.73*				
7 Impression Management	.29*	.50*	-.04	-.02	.09	-.08			
8 Behavioral Interview Perf	.29*	.15	.26*	.60*	.64*	.64*	.00		
9 Situational Interview Perf	.36*	.12	.36*	.61*	.56*	.57*	.07	.55*	
Mean	3.64	3.90	21.69	3.02	3.35	3.22	3.48	4.03	4.2
SD	.46	.44	6.76	.67	.54	.53	.49	.91	.94
$\alpha$	.77	.87		.88	.83	.86	.64	.93	.94
ICC				.84	.79	.64		.82	.80

$N = 81$ ;  $^{\wedge} = p < .10$ ,  $* p < .05$

ICC intraclass correlation coefficient (C, 2) between the ratings of the three judges, ST Selection Test, EA Evaluator Assessment

Correlations underlined and in italics represent correlations between respective ST and EA traits

Correlations in the upper right diagonal represent different source correlations across raters

To measure evaluator ratings of GMA, two items were taken from the scale developed by Gignac et al. (2004). The items include “intelligent, bright” and “has a good vocabulary.”

Ratings for each of the five interview performance dimensions were rated on a 7-point scale with behavioral anchors based on DeGroot and Motowidlo (1999). An example of a leadership anchor is “Little or no effort to seek out opportunities for leadership/low/1” to “Accepts leadership roles when opportunities arise/moderate/4” to “Gravitates naturally to leadership positions/high/7.” The full measures are reported in the “Appendix” section.

## Results

Table 1 reports descriptive statistics and correlations. All coefficient alphas were satisfactory, ranging from .77 for selection test extraversion to .94 for situational interview performance, with the exception of the impression management control variable ( $\alpha = .64$ ). In addition, interrater reliabilities were estimated using ICC (2,3) (two-way random average measures intraclass correlation coefficients using 3 raters; see Shrout and Fleiss 1979 for a review) resulting in .84 for extraversion, .79 for EI, .64 for GMA, .82 for behavioral interview performance, and .80 for situational interview performance. These values are somewhat higher than values found through meta-analysis of other ratings (Connolly et al. 2007) which are highest for extraversion (.66). Thus, as reliability is a necessary condition for validity, GMA, EI, and extraversion appear to show promise as predictors of interview performance, such that it is measured reliably and are rated consistently by evaluators.

Table 1 also forms a multi-trait, multi-method matrix (MTMM; Campbell and Fiske 1959) with extraversion, EI, and GMA serving as traits crossed with selection test and evaluator-assessed methods. Inferences about convergent and discriminant validity are made by analyzing the patterns of correlations in the MTMM matrix. Convergent validity is inferred if the correlations on the validity diagonal are significantly greater than zero and are sufficiently large to warrant further investigation. All correlations along the validity diagonal (indicated in italics and underlined) were significant, therefore, providing evidence of convergent validity. Discriminant validity, then, is assessed by evaluating relationships among three components of each matrix: (1) the validity diagonal indicates correlations between a measure and itself across methods, (2) the different-trait/different-method values (above and below the validity diagonals in the rectangle) contain correlations between a measure in one condition (selection test) and other measures in the other condition (evaluator assessments), and (3) the different-trait/same-method triangles (set off with solid lines) include correlations among the various measures within one condition.

A comparison of the validity diagonal with the different-trait/different-method values indicates a slightly larger correlation between selection test extraversion and evaluator assessment EI (.28) than same-trait correlations for EI (.22) and GMA (.27). When evaluating the different-trait/same-method triangles for selection tests, the correlation between extraversion and EI is relatively large (.40). When evaluating the different-trait/same-method triangles for evaluator assessments, the intercorrelations between the evaluator-rated traits of GMA and EI (.73) are higher than the correlations between evaluator-rated traits and

interview performance (ranging from .56 to .64). Whereas the evaluator-rated traits include same source correlations in which the same evaluator rated each trait, the correlations between evaluator trait ratings and interview performance are all based on the trait and performance ratings coming from different sources. To address this same source issue, different source trait intercorrelations were calculated in an effort to minimize common-method variance (Podsakoff et al. 2003). This was done for each trait by correlating (a) ratings of a trait from rater 1 with the average of the two remaining trait ratings from raters 2 and 3, (b) ratings of a trait from rater 2 with the average of the two remaining trait ratings from raters 1 and 3, and (c) ratings of a trait from rater 3 with the average of the two remaining trait ratings from raters 1 and 2. These three sets of correlations were then averaged (using Fisher's  $z$  transformations) to produce different source correlations. A comparison of the different source evaluator trait correlations (ranging from .36 to .50) is all lower than the correlations between evaluator traits and interview performance (ranging from .56 to .64). Taken together, the MTMM analysis generally yields discriminant validity across traits and methods, but with some exceptions.

Correlations between selection tests and evaluator assessments were statistically significant but modest in magnitude for extraversion ( $r = .36, p < .001$ ), EI ( $r = .22, p = .048$ ), and GMA ( $r = .27, p = .015$ ). These results are in line with meta-analytic results regarding self- and other-assessed traits (Connolly et al. 2007), and support our argument that tests scores and evaluator assessments are in part tapping the same construct, yet are distinct enough to allow for unique prediction for both measurement approaches. Impression management is significantly correlated with selection test extraversion and EI ( $r = .29, p = .008$  and  $r = .50, p < .001$ , respectively), but not with the test of GMA or any trait assessments or interview performance evaluations.

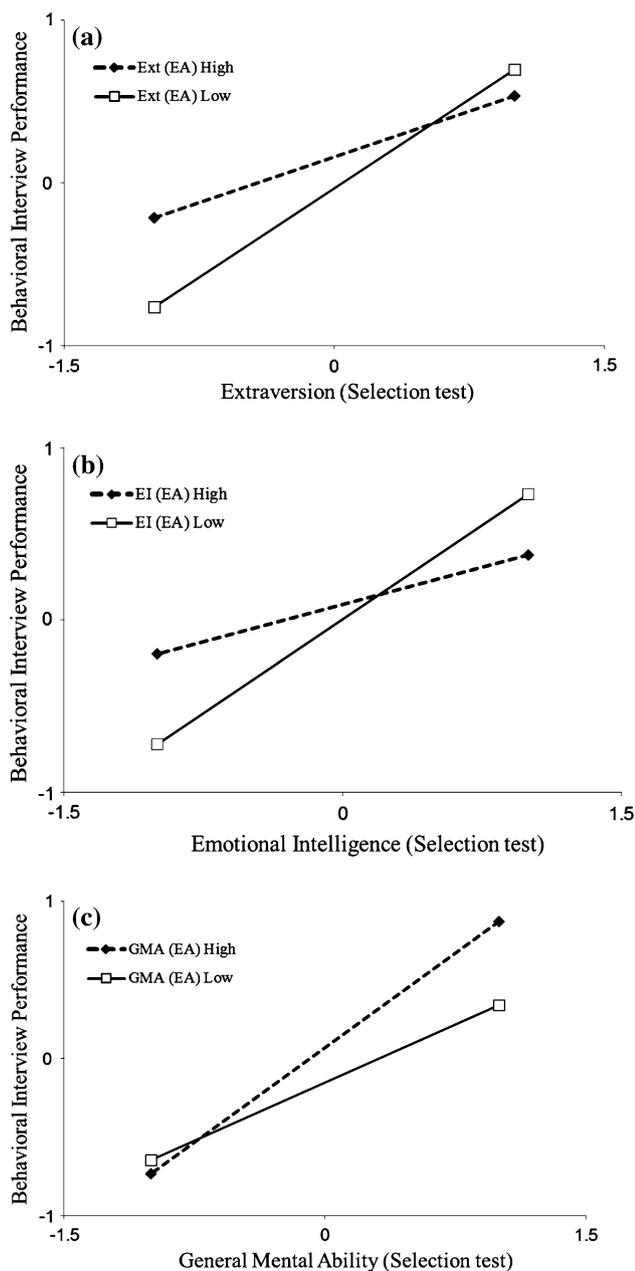
The prediction of behavioral interview performance was examined for selection tests and evaluator assessments of GMA, EI, and extraversion; extraversion ( $r = .29, p = .008$  and  $.60, p < .001$ , respectively), EI ( $r = .15, n.s.$  and  $.64, p < .001$ , respectively), and GMA ( $r = .26, p = .018$  and  $r = .64, p < .001$ , respectively). Thus, these results provide initial support for hypotheses 1a, 1b, 2b, 3a, and 3b, but not for 2a consisting of selection tests for emotional intelligence. The same behavioral interview was assessed for both trait evaluator ratings and interview performance, which creates an unfair advantage for evaluator assessments when comparing the magnitude of correlations with interview performance. In other words, both sets of evaluators are watching the exact same behaviors when making assessments. As noted earlier, to address this limitation, we had an independent set of three evaluators (not used in the other

assessments) rate interview performance on only the situational question set embedded within the full interview. As such, the situational interviews represent an entirely independent set of behaviors stemming from an entirely independent set of interview questions evaluated by different sets of raters. Results indicate a similar pattern of correlations across the two different assessment sources (self versus other assessments): extraversion ( $r = .36, p < .001$  and  $.61, p < .001$ , respectively), EI ( $r = .12, n.s.$  and  $.56, p < .001$ , respectively), and GMA ( $r = .36, p < .001$  and  $r = .57, p < .001$ , respectively). These results further support hypotheses 1a, 1b, 2b, 3a, and 3b, but not hypothesis 2a for emotional intelligence.

Hypothesis 4a posited that GMA would be the strongest predictor among the selection tests. Results demonstrate that extraversion and GMA are relatively equivalent predictors of interview performance, though stronger than the results found for EI. Thus, Hypothesis 4a is not supported. For Hypothesis 4b, we proposed that GMA would be the strongest predictor among evaluator-rated traits. Across behavioral and situational interviews, GMA, EI, and extraversion were relatively equivalent and not statistically different from one another. Thus, Hypothesis 4b is not supported. For Hypothesis 4c, we posited that evaluator assessments would be stronger predictors of interview performance than selection tests. Across behavioral and situational interviews, selection test correlations ranged from .12 to .36, while evaluator assessment correlations ranged from .56 to .64. Thus, Hypothesis 4c is fully supported.

Hypotheses 5a, 5b, and 5c posited that the relationship between each selection test traits of GMA, EI, and extraversion would interact with its respective evaluator-assessed trait to predict interview performance, such that low levels of both self-ratings and assessor ratings would yield the lowest levels of interview performance. We ran a three-stage hierarchical regression with impression management in the first stage, the respective selection test and evaluator traits added to the second stage, and the interaction term added to the third stage. As suggested by Bing et al. (2007), to increase statistical power, we use a one-tailed test for our interactions because they were predicted a priori and statistical power for detecting interactions in field research is all too often overly low (Chaplin 1991; Morris et al. 1986). All statistically significant interactions were graphed by first standardizing all variables, obtaining the standardized betas in regression, then plotting the interactions at one standard deviation above and below the mean using an Excel interaction macro.

For GMA, the interaction was significant for both the behavioral ( $\Delta R^2 = 2\%, p = .047$ ) and situational interviews ( $\Delta R^2 = 4\%, p = .012$ ). However, as shown in Fig. 1, the pattern of results is not completely in line with our theorizing. We hypothesized that low levels of both



**Fig. 1** The interactions of selection test and evaluator assessment (EA) in the prediction of behavioral interview performance for **a** extraversion and **b** emotional intelligence, and **c** general mental ability

selection test and evaluator assessments will yield the lowest levels of interview performance. While this is what was found for GMA, the graph also shows that interview performance is low when selection test GMA is low, but evaluator-assessed GMA is high. Further, interview performance was maximized when selection test and the evaluator assessments of GMA were high. As such, Hypothesis 5a is mostly supported. As shown in Table 2, a significant interaction was found for EI in the behavioral interview ( $\Delta R^2 = 4\%$ ,  $p = .021$ ) and in the situational

interview ( $\Delta R^2 = 3\%$ ,  $p = .037$ ). As shown in Fig. 1, again, the lowest level of interview performance was obtained when both selection test and evaluator ratings of EI were low. Thus, Hypothesis 5b is fully supported. Similarly, as shown in Table 2, a significant interaction was found for extraversion in the behavioral interview ( $\Delta R^2 = 3\%$ ,  $p = .024$ ), but not in the situational interview. As shown in Fig. 1, the lowest level of interview performance was obtained when both selection test and evaluator ratings of extraversion were low. Thus, Hypothesis 5c is partially supported. Finally, we also chose to report the total variance in interview performance that is accounted for with all 6 measures and the three interactive effects combined. Results indicate an uncorrected multiple  $r$  of .71 ( $R^2 = 50.2\%$ ) for the behavioral interview and a multiple  $r$  of .71 ( $R^2 = 50.4\%$ ) for the situational interview (Fig. 2).

## Discussion

This study is the first to compare the relative impact of traits assessed via selection test versus evaluator-assessed characteristics on interview performance. Our results indicate that extraversion, GMA, and EI have an important influence on interview performance, but this effect differs when these traits are measured via selection test versus evaluator assessment. Because of the quality of trait information yielded in interviews, organizational representatives may indeed garner unmeasured information about applicants, which holds the potential to be useful for selection and other important organizational functions such as training and organizational fit. At a minimum, these judgments influence the ratings of job-relevant qualifications in structured behavioral and situational interviews. Further, traits measured via a selection test generally interact with their other-assessed counterpart in predicting interview performance.

The results of this study provide insight into how interviewers utilize the information that is produced in an employment interview. Specifically, we tested and found support for the impact of traits influencing interview performance in a structured interview, showing that extraversion, EI, and GMA have substantial influence on interview performance, though the level of influence depends on whether the traits are measured as traditional selection tests or evaluator assessments. We found little difference in the magnitude of the correlations between the traits of GMA, EI, and extraversion and interview performance when comparing magnitudes for selection tests and for evaluator assessments independently. However, evaluator assessments of these traits are stronger predictors of interview performance than are selection tests.

**Table 2** Hierarchical regression results

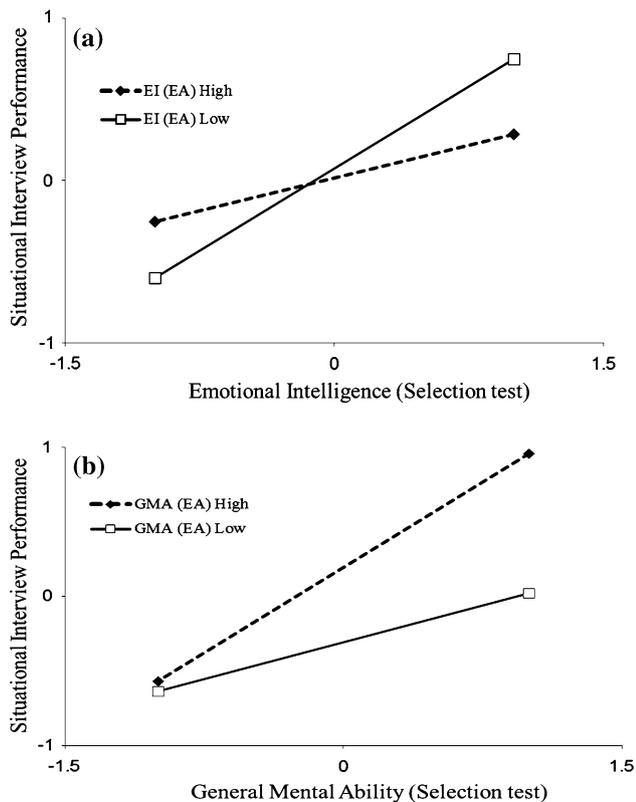
	Extraversion				Emotional intelligence				General mental ability			
	$\beta$	$R$	$\Delta R^2$	$\Delta F$	$\beta$	$R$	$\Delta R^2$	$\Delta F$	$\beta$	$R$	$\Delta R^2$	$\Delta F$
Behavioral interview												
Step 1		.00	.00	.00		.00	.00	.00		.00	.00	.00
Impression management	.00				.00				.00			
Step 2		.60*	.36*	22.25*		.57*	.33*	19.27*		.65*	.42*	28.56*
Impression management	-.02				-.09				.05			
Selection test	.09				.07				.10			
Evaluator assessment	.56*				.56*				.62*			
Step 3		.63*	.03*	4.07*		.60*	.04*	4.32*		.66*	.02*	2.89*
Impression management	-.04				-.05				.05			
Selection test	.10				.04				.11			
Evaluator assessment	.55*				.51*				.65*			
Interaction	-.18*				-.20*				.15*			
Situational interview												
Step 1		.07	.00	.36		.07	.00	.36		.07	.00	.36
Impression management	.07				.07				.07			
Step 2		.63*	.40*	26.18*		.52*	.27*	14.54*		.61*	.37*	23.05*
Impression management	.04				.02				.11			
Selection test	.15				-.01				.23*			
Evaluator assessment	.56*				.52*				.50*			
Step 3		.64	.01	.83		.55*	.03*	3.30*		.64*	.04*	5.41*
Impression management	.03				.06				.12			
Selection test	.15				-.03				.25*			
Evaluator assessment	.56*				.47*				.55*			
Interaction	-.08				-.18*				.21*			

$N = 81$ ;  $\wedge = p < .10$ , \*  $p < .05$

Our research is also the first to assess the potential interactive effects of selection test and evaluator ratings of individual differences for the prediction of interview performance. Using socioanalytic theory (Hogan 1996, 2007), we were able to demonstrate that the identity and reputation aspects of these constructs combine to further predict interview performance. More specifically, low levels on both measurement approaches for extraversion and EI lead to the lowest levels of interview performance. For GMA, high levels of both measurement approaches lead to the highest levels of interview performance. However, the form of the interactions for GMA does not match the form we proposed a priori. Specifically, as was found for extraversion and EI, we proposed that low levels of both measurement approaches would yield the lowest levels of interview performance. For GMA, we found that those low in GMA via selection test performed equally poorly whether or not evaluator-assessed GMA was high or low. However, interviewees performed best in both behavioral and situational interviews when high in both GMA via selection test and high in GMA via evaluator assessment.

One possible reason for these differences in the form of interactive effects may be due to the differences in measurement between these traits. Specifically, extraversion and EI are assessed using self-assessed selection tests, while GMA is measured via a performance test with right and wrong answers. Another reason why we see differences in our results between personality traits and cognitive ability is that whereas they are both individual differences, they represent different aspects of the individual that have different effects on performance. The absence of a trait or ability is generally viewed as being detrimental, and the presence of a trait or ability is viewed as being beneficial—no matter how these variables are measured (i.e., selection tests or evaluator ratings). Future research should investigate these divergent interactive effects.

One aspect of traditional assessment methods and job interviews that may have affected our results is the issue of social desirability/faking. Traditional self-assessed trait measures, such as EI (Klumper 2008) and extraversion (McFarland and Ryan 2000), have been found to be influenced by socially desirable responding, while tests of



**Fig. 2** The interactions of selection test and evaluator assessment (EA) in the prediction of situational interview performance for **a** emotional intelligence and **b** general mental ability

GMA have been shown to be resistant to such distortions (Ones et al. 1996). The effects of socially desirable responding on trait ratings in the interview context, however, are less clear. Self-presentation tactics (Barrick et al. 2009) and impression management tactics (Kristof-Brown et al. 2002) are shown to impact interview performance, while self-presentation may represent a systematic source of inaccuracy in interview performance (Van Iddekinge et al. 2005; Posthuma et al. 2002). However, Van Iddekinge and colleagues found that interview ratings of traits did not significantly differ in factor structure or mean differences across honest and applicant faking groups. Our findings support this existing research, such that selection test extraversion and EI correlate significantly with impression management, while evaluator-rated traits and GMA test scores do not relate to impression management. However, the impact of response distortion is known to be stronger for job applicants than job incumbents (Rosse et al. 1998). As such, future research should evaluate the impact of social desirability on evaluator assessments of traits with actual job applicants.

Our research demonstrates that different methods of assessing individual differences yield unique information

about aspects of these traits. For example, whereas a modest correlation exists between evaluator-assessed GMA and selection test GMA, we believe this highlights the different facets of GMA that are being assessed with these different methods. Thus, the lack of strong overlap between the two can be viewed as a strength of multiple methods of assessing GMA and its explanation of subsequent interview performance. Alternatively, as pointed out by a reviewer, an interviewee's style of speech using complex vocabulary, diction, and sentence structure could provide clues to the degree of GMA, but the content of the information provided may not be relevant to the question posed during the interview. This failure to address the inquiry with adequate information (though skillfully phrased) would result in a lower performance ranking and muddy the measurement and impact of the interviewee's actual intelligence. The degree to which this may be the case is an area that could be explored in future research efforts. By examining our variables through MTMM techniques, we were able also to establish both convergent and discriminant validity (with a few exceptions) among our measurements of the traits. This further establishes our position that these measures are tapping unique aspects of the traits that impact interview performance.

From a practical perspective, the present research has important implications as we attempt to understand the dynamics of the interview more fully. Although the current study does not assess the criterion-related validity of interview-rated traits on job performance, it does help establish that certain job-relevant traits can both be reliably assessed in a structured interview and relate to interview performance. While hiring agents sometimes choose to not use selection tests due to concerns over faking and the perceived irrelevance of such measures, they typically feel compelled to engage in interviewing. This study suggests that interviews may serve as a valuable means of assessing traits that often go unmeasured in spite of their potential benefit. Our findings suggest that interviewers—perhaps subconsciously—assess such constructs as GMA and personality, which then factors into their ratings of overall interview performance. The evidence suggests that this occurs even in a structured approach which is intended to focus interviewer judgments on clear job-related requirements and away from more subjective and general assessments. Perhaps structured interviews produce even greater potential for assessments of this nature than less targeted interviews. The data in this study suggest that interviewing may be a valuable source of information regarding GMA and personality which provides a rationale for why interviewers are seemingly reluctant to give up face-to-face interviewing in spite of evidence that there are more valid selection instruments.

## Limitations and Future Research

Although the present study offers a novel perspective on the drivers of ratings of interview performance, we are mindful of certain limitations that may have affected our results. First, interviewees and evaluators completed different measures of the constructs. Further, although our evaluator assessments obtained sufficient internal consistency reliability, these constructs were measured with only two items per construct, whereas our selection tests consisted of 12–50 items. As such, the relationships observed in the current study may be due to traits, to differences in measurement, or both. In addition, evaluator ratings and interview performance were both generated from the same narrow method (the video-recorded job interviews), thus likely introducing common-method bias. In this vein, future research should include evaluator assessments outside of the interview context, such as co-worker ratings of extraversion, EI, and GMA to investigate how these assessments compare their predictive validities to traditional selection tests. A second potential limitation is that the results in the current study may not generalize to less-structured interviews and even-structured interviews for jobs beyond the focal position used in our study. Third, the current study did not include ratings of job performance. Rather, the dependent variable in this study was interview performance. Future research should identify the extent that evaluator-assessed predictors of interview performance translate into job performance. Further, because a variety of the traits that predict interview performance ratings in the current study have been found to be more predictive of job performance in some jobs than in others, future research should assess trait relevance to the job as a key moderator of the relationship between interview ratings and job performance. For example, perhaps structured interviews are more valid for jobs requiring higher levels of extraversion due to the degree of influence that this trait has on interview ratings.

The focus of the current study was to further elucidate the inadvertent impact of traits on structured interviews, interviews typically designed to assess potential job performance. In this vein, the job-relevant questions designed for the job of Youth Treatment Specialist may be more salient for some traits than others. As such, this raises issues with respect to the generalizability of our results across interviews developed for different types of jobs. Further, emerging research (e.g., Van Iddekinge et al. 2005; Roth et al. 2005) has sought to design interviews specifically to assess personality traits using structured interviews, while other research (Blackman 2002; Blackman and Funder, 2002; Townsend et al. 2007) has sought to assess personality traits using unstructured interviews. Future research should investigate the relative value of

these divergent approaches. A related aspect in design and application is the length of the interview. Longer interviews should yield more information and, thus, more accurate trait assessments. It would also be of interest to assess whether approaches to interviewing that have been developed with the intention of reducing the potential for subjective bias (such as written answers to questions) influence the assessment of GMA and personality, thus impacting the potential of such approaches to provide valuable information that may lead to greater selection validity.

## Conclusion

Our research provides impetus for continued examination of the extent to which interviewers utilize the information produced through interviews to form judgments about applicant traits and how those judgments influence the evaluation of interview performance. This study finds that the traits of GMA, EI, and extraversion are relatively equivalent predictors of interview performance. In addition, evaluator assessments of these traits are stronger relative predictors of interview performance. Finally, we establish that selection test and evaluator assessments of each trait interact to explain more variance in interview performance. This demonstrates that multiple measures of a trait can provide unique and relevant explanations of interview performance.

**Acknowledgments** We would like to thank Tim DeGroot and Lilli A. Bishop for helpful comments.

## Appendix

### Behavioral Interview Questions

Tell me about a time when you were challenged to get somebody to do something they really did not want to do.

1. Please describe a time when you had to work with someone who was difficult to get along with.
2. At times we are put in situations where we find ourselves correcting someone's behavior because it is inappropriate, offensive, or just plain wrong for other reasons. Tell me about a situation where you had to confront someone who was doing something wrong.
3. Please explain something you have done in a work situation that shows how creative or innovative you can be.
4. Please explain a recent decision that you had to make that was particularly challenging or complicated.

### Situational Interview Questions

1. Imagine you were working with a fellow worker whom you knew greatly disliked performing a particular job task. You were in a situation where you needed this task completed, and this employee was the only one available to assist you. What would you do to motivate the employee to perform this task?
2. Imagine being in a situation with both a co-worker and a client in which the client is being unreasonable. Your co-worker appears frustrated and begins to make comments that may be construed as sarcastic and offensive. How do you handle this situation?
3. Imagine that, as part of a living unit activity, you are dealing with client who is difficult to deal with. In front of all the other clients, they refuse to follow your directives. How would you go about dealing with the client?
4. Suppose you are working on an important report and become increasingly uncertain whether or not you will complete the project by the stated deadline set for you by your supervisor. How would you deal with this situation?
5. Imagine that your boss is sick and you are asked to fill in for him or her for a few weeks. One particular task requiring attention is to plan for an upcoming outing to a local park including yourself, two additional staff, and 12 clients. What would you do to prepare for this outing?

Behaviorally Anchored Rating Scales

**Leadership:** taking charge, giving direction, delegating and following up, and motivating others.

\_\_\_ 7 High – Naturally gravitates toward leadership positions.  
Actively pursues opportunities to direct others.  
Seeks opportunities to direct and motivate others to accomplish group goals.

\_\_\_ 6

\_\_\_ 5

\_\_\_ 4 Mod – Accepts leadership roles when given the opportunity.  
Directs and motivates others to achieve group goals.  
Delegates and follows up.

\_\_\_ 3

\_\_\_ 2

\_\_\_ 1 Low - Expresses little or no effort to seek out opportunities for leadership.  
Reluctant to accept leadership roles when given the opportunity.  
Does not delegate or follow up.

**Initiative:** Working hard to reach goals and meet deadlines, persisting to overcome obstacles, taking initiative, producing innovative and workable solutions to problems.

\_\_\_ 7 High – Works as much as it takes to overcome obstacles and complete all aspects of the task on time, leaving no loose ends.  
Demonstrates unusual resourcefulness in devising imaginative and workable solutions to difficult problems.  
Takes initiative to develop and implement new solutions or new courses of action.

\_\_\_ 6

\_\_\_ 5

Works hard to overcome obstacles and complete most aspects of the task on time without sacrificing important details.  
\_\_\_ 4 Mod – Finds workable solutions to problems and eventually finds a way around obstacles.  
Develops and recommends new solutions or new courses of action when requested.

\_\_\_ 3

\_\_\_ 2

Extends deadlines or unnecessarily asks for help when encountering difficult obstacles.  
Shows no evidence of having grappled successfully with difficult problems or overcoming major obstacles.  
\_\_\_ 1 Low - Refuses or fails to develop new solutions or new courses of action when requested.

**Persuasiveness:** Persuading others to accept own ideas, showing consideration for others' feelings when disagreeing with them, confronting others assertively yet tactfully when necessary.

\_\_\_ 7 High – Successfully defends own point of view despite strong objections from others.  
Presents contrary positions sensitively without offending others or damaging their self-esteem.  
Confronts others assertively yet tactfully when appropriate to correct their behavior.

\_\_\_ 6

\_\_\_ 5

\_\_\_ 4 Mod – Tries to defend own point of view when appropriate and usually succeeds.  
Tries to present contrary positions sensitively but may inadvertently offend others or damage their self-esteem.  
Tries to be tactful when confronting others to correct their behavior, but may inadvertently offend them.

\_\_\_ 3

\_\_\_ 2

Lacks confidence and forcefulness, or...  
Presents own point of view in a domineering, arrogant, or condescending way.  
Avoids presenting contrary positions, or...  
Presents contrary positions insensitively with no effort to avoid offending others.  
Avoids confronting others to correct their behavior, or ...  
\_\_\_ 1 Low - Confronts others aggressively without trying to be tactful.

**Thoroughness:** Gathering as much information about a problem as possible, attending carefully to details, generating alternative solutions and evaluating them thoroughly.

\_\_\_ 7 High - Examines and resolves all aspects of a problem.  
Keeps track of all relevant details and leaves no loose ends.  
Uses all available time to gather as much information as possible about an issue.

\_\_\_ 6

\_\_\_ 5

\_\_\_ 4 Mod - Resolves key aspects of a problem sufficiently, especially the readily apparent aspects.  
Keeps track of many important details but may miss some less obvious but relevant details.  
Gathers as much information as necessary to develop a satisfactory solution.

\_\_\_ 3

\_\_\_ 2

Overlooks key aspects of problems.  
Fails to recognize that more information is needed and/or potentially available.  
\_\_\_ 1 Low - Tries to get by on only vague information or untested and unsupported assumptions.

**Oral Communication:** Focusing on relevant issues, presenting information clearly and concisely, speaking audibly and fluently, using appropriate grammar, using animated gestures and maintaining eye contact appropriately.

\_\_\_ 7 High - Volunteers relevant information clearly, concisely, and in an interesting way that captures listeners' attention.

Maintains effective eye contact and tone of voice to help convey the message.

Speaks fluently and articulately using effective gestures to help convey the message.

\_\_\_ 6

\_\_\_ 5

\_\_\_ 4 Mod - Answers questions adequately with relevant information.

Uses moderate eye contact and speaks at an appropriate tone of voice.

Speaks clearly with appropriate grammar and body language.

\_\_\_ 3

\_\_\_ 2

Answers questions in a disjointed and confusing way with irrelevant information.

Does not make eye contact or does not use an appropriate tone of voice.

\_\_\_ 1 Low - Is unclear or awkward when speaking.

## References

- Ashkanasy, N. M., Hartel, C. E. J., & Daus, C. S. (2002). Diversity and emotion: The new frontiers in organizational behavior research. *Journal of Management*, *38*, 307–338.
- Baron, R. A. (1993). Interviews moods and evaluations of job applicant: The role of applicant qualifications. *Journal of Applied Social Psychology*, *23*, 253–271.
- Barrick, M. R., Shaffer, J. A., & DeGrassi, S. W. (2009). What you see may not be what you get: Relationships among self-presentation tactics and ratings of interview and job performance. *Journal of Applied Psychology*, *94*, 1394–1411.
- Beatty, R. N. (1986). *The five-minute interview*. New York: Wiley.
- Berry, C. M., Sackett, P. R., & Landers, R. N. (2007). Revisiting interview-cognitive ability relationships: Attending to specific range restriction mechanisms in meta-analysis. *Personnel Psychology*, *60*, 837–874.
- Bing, M. N., LeBreton, J. M., Davison, H. K., Migetz, D. Z., & James, L. R. (2007). Integrating implicit and explicit social cognitions for enhanced personality assessment: A general framework for choosing measurement and statistical methods. *Organizational Research Methods*, *10*(2), 346–389.
- Blackman, M. C. (2002). Personality judgment and the utility of the unstructured employment interview. *Basic and Applied Social Psychology*, *24*, 241–250.
- Blackman, M. C., & Funder, D. C. (2002). Effective interview practices for accurately assessing counterproductive traits. *International Journal of Selection and Assessment*, *10*, 109–116.
- Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and

- intelligence. *Journal of Personality and Social Psychology*, 86, 599–614.
- Brody, N. (1992). *Intelligence* (2nd ed.). San Diego, CA: Academic Press.
- Burnett, J. R., Fan, C., Motowidlo, S. J., & DeGroot, T. (1998). Interview notes and validity. *Personnel Psychology*, 51, 375–396.
- Burroughs, J. W., Drew, D. R., & Hallman, W. K. (1991). Predicting personality from personal possessions: A self-presentational analysis. *Journal of Social Behavior and Personality*, 6, 147–163.
- Caldwell, D. F., & Burger, J. M. (1998). Personality characteristics of job applicants and success in screening interviews. *Personnel Psychology*, 51, 119–136.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.
- Campion, M. A., Palmer, D. K., & Campion, J. E. (1997). A review of structure in the selection interview. *Personnel Psychology*, 50, 655–702.
- Campion, M. A., Pursell, E. D., & Brown, B. K. (1988). Structured interviewing: Raising the psychometric properties of the employment interview. *Personnel Psychology*, 41, 25–42.
- Chaplin, W. F. (1991). The next generation of moderator research in personality psychology. *Journal of Personality*, 59, 143–178.
- Choi, S., & Kluemper, D. H. (2011). The relative utility of differing measures of emotional intelligence: Other-rated EI as a predictor of social functioning. *European Review of Applied Psychology*, 93, 270–277.
- Choi, S., Kluemper, D. H., & Sauley, K. S. (2011). What if we fake emotional intelligence? A test of criterion validity attenuation. *Journal of Personality Assessment*, 93, 270–277.
- Cole, M. S., Feild, H. S., & Giles, W. F. (2003). Using recruiter assessments of applicants resume content to predict applicant mental ability and Big Five personality dimensions. *International Journal of Selection and Assessment*, 11, 78–88.
- Cole, M. S., Feild, H. S., & Stafford, J. O. (2005). Validity of resume reviewers inferences concerning applicant personality based on resume evaluation. *International Journal of Selection and Assessment*, 13, 321–324.
- Connolly, J. J., Kavanagh, E. J., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International Journal of Selection and Assessment*, 15, 110–117.
- Conway, J. M., Jako, R. A., & Goodman, D. F. (1995). A meta-analysis of interrater and internal consistency reliability of selection interviews. *Journal of Applied Psychology*, 80, 565–579.
- Cortina, J. M., Golstein, N. B., Payne, S. S., Davison, H. K., & Gilliland, S. W. (2000). The incremental validity of interview scores over and above cognitive ability and conscientiousness scores. *Personnel Psychology*, 53, 325–351.
- Costa, P. T., & McCrae, R. R. (1985). *The NEO personality inventory manual*. Odessa, FL: Psychological Assessment Resources Inc.
- DeGroot, T., & Gooty, J. (2009). Can nonverbal cues be used to make meaningful personality attributions in employment interviews? *Journal of Business and Psychology*, 24, 179–192.
- DeGroot, T., & Motowidlo, S. (1999). Why visual and vocal interview cues can affect interviewers' judgments and predict job performance. *Journal of Applied Psychology*, 84, 986–993.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417–440.
- Dipboye, R. L. (1992). *Selection interviews: Process perspectives*. Cincinnati, OH: South-Western.
- Dipboye, R. L. (1994). Structured and unstructured selection interviews: Beyond the job-fit model. In G. R. Ferris (Ed.), *Research in personnel and human resources management* (Vol. 12). Greenwich, CT: JAI Press.
- Dunn, W. S., Mount, M. K., Barrick, M. R., & Ones, D. S. (1995). Relative importance of personality and general mental ability in managers' judgments of applicant qualifications. *Journal of Applied Psychology*, 80, 500–509.
- Fast, L. A., & Funder, D. C. (2008). Personality as manifest in word use: Correlations with self-report, acquaintance report, and behavior. *Journal of Personality and Social Psychology*, 94, 334–346.
- Fisher, C. D., & Ashkanasy, N. M. (2000). The emerging role of emotions in work life: An introduction. *Journal of Organizational Behavior*, 21, 123–129.
- Fox, S., & Spector, P. E. (2000). Relations of emotional intelligence, practical intelligence, general intelligence, and trait affectivity with interview outcomes: It's not all just "G". *Journal of Organizational Behavior*, 21, 203–220.
- Funder, D. C., & Sneed, C. D. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology*, 64, 479–490.
- Gignac, G. E., Stough, C., & Loukomitis, S. (2004). Openness, intelligence, and self-report intelligence. *Intelligence*, 32, 133–143.
- Gill, A. J., Oberlander, J., & Austin, E. (2006). Rating e-mail personality at zero acquaintance. *Personality and Individual Differences*, 40, 497–507.
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B. Jr. (2003). A very brief measure of the Big Five personality domains. *Journal of Research in Personality*, 37, 504–528.
- Hakel, M. D. (1982). Employment interviewing. In K. M. Rowland & G. R. Ferris (Eds.), *Personnel Management* (pp. 129–155). Boston: Allyn and Bacon.
- Highhouse, S., & Bottrill, K. V. (1995). The influence of social (mis)information on memory for behavior in an employment interview. *Organizational Behavior and Human Decision Processes*, 62, 220–229.
- Hofstee, W. K. (1994). Who should own the definition of personality? *European Journal of Personality*, 8, 149–162.
- Hogan, R. T. (1996). A socioanalytic perspective on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality* (pp. 163–179). New York: Guilford Press.
- Hogan, R. T. (2007). *Personality and the fate of organizations*. Mahwah, NJ: Erlbaum.
- Hogan, R. T., & Shelton, D. S. (1998). A socioanalytic perspective on job performance. *Human Performance*, 11, 129–144.
- Huffcutt, A. I., & Arthur, W. J. (1994). Hunter and Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology*, 79, 577–580.
- Huffcutt, A. I., Conway, J. M., Roth, P. L., & Stone, N. J. (2001a). Identification and meta-analytic assessment of psychological constructs measured in employment interviews. *Journal of Applied Psychology*, 86, 897–913.
- Huffcutt, A. I., & Roth, P. L. (1998). Racial group differences in interview evaluations. *Journal of Applied Psychology*, 83, 288–297.
- Huffcutt, A. I., Roth, P. L., & McDaniel, M. A. (1996). A meta-analytic investigation of cognitive ability in employment interview evaluations: Moderating characteristics and implications for incremental validity. *Journal of Applied Psychology*, 81, 459–473.
- Huffcutt, A. I., Van Iddekinge, C. H., & Roth, P. L. (2011). Understanding applicant behavior in employment interviews: A theoretical model of interviewee performance. *Human Resource Management Review*, 21, 353–367.
- Huffcutt, A. I., Weekley, J. A., Wiesner, W. H., DeGroot, T., & Jones, C. (2001b). Comparison of situational and behavior description

- interview questions for higher-level positions. *Personnel Psychology*, 54, 619–644.
- Hunter, J. E., & Hirsch, H. R. (1987). Applications of meta-analysis. In C. I. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (Vol. 2, pp. 257–321). New York: Wiley.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72–98.
- Hunter, J. E., Schmidt, F. L., & Le, H. (2006). Implications of direct and indirect range restriction for meta-analysis methods and findings. *Journal of Applied Psychology*, 91, 594–612.
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The five revisited. *Journal of Applied Psychology*, 85, 869–879.
- Janz, T. (1982). Initial comparisons of patterned behavioral description interviews versus unstructured interviews. *Journal of Applied Psychology*, 67, 577–580.
- John, O. P., & Srivastava, S. (1999). The “Big Five” trait taxonomy: History, measurement, and theoretical perspectives. In L. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). New York: Guilford.
- Kluemper, D. (2008). Trait emotional intelligence: The impact of core-self evaluations and social desirability. *Personality and Individual Differences*, 44, 1402–1412.
- Kluemper, D., DeGroot, T., & Choi, S. (2013). Emotion management ability: Predicting task performance, citizenship, and deviance. *Journal of Management*, 39, 878–905.
- Kluemper, D. H., McLarty, B. D., & Bing, M. N. (2014). Acquaintance ratings of the Big Five personality traits: Incremental validity beyond and interactive effects with self-reports in the prediction of workplace deviance. *Journal of Applied Psychology*. doi:10.1037/a0037810
- Kluemper, D. H., & Rosen, P. (2009). Future employment selection methods: Evaluating social networking websites. *Journal of Managerial Psychology*, 24, 567–580.
- Kluemper, D., Rosen, P., & Mossholder, K. (2012). Social networking websites, personality ratings, and the organizational context: More than meets the eye? *Journal of Applied Social Psychology*, 42, 1143–1172.
- Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of Personality*, 64, 311–337.
- Konig, C. J., Melchers, K. G., Kleinmann, M., Richter, G. M., & Klehe, U. C. (2007). Candidates’ ability to identify criteria in nontransparent selection procedures: Evidence from an assessment center and a structured interview. *International Journal of Selection and Assessment*, 15, 283–292.
- Kristof-Brown, A. L., Barrick, M. R., & Franke, M. (2002). Influences and outcomes of candidate impression management in job interviews. *Journal of Management*, 28, 27–46.
- Latham, G. P., Saari, L. M., Pursell, E. D., & Campion, M. A. (1980). The situational interview. *Journal of Applied Psychology*, 65, 422–427.
- Lievens, F., Highhouse, S., & DeCorte, W. (2005). The importance of traits and abilities in supervisors’ hireability decisions as a function of method of assessment. *Journal of Occupational and Organizational Psychology*, 78, 453–470.
- Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence. In P. Salovey & D. J. Sluyter (Eds.), *Emotional development and emotional intelligence*. New York: Basic Books.
- McCrae, R. R., & Costa, P. T. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research*. New York: Guilford Press.
- McCrae, R. R., & Weiss, A. (2007). Observer ratings of personality. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology*. New York: Guilford Press.
- McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, 79, 599–616.
- McFarland, L. A., & Ryan, A. M. (2000). Variance in Faking Across Noncognitive Measures. *Journal of Applied Psychology*, 85, 812–821.
- Morris, J. H., Sherman, J., & Mansfield, E. R. (1986). Failures to detect moderating effects with ordinary least squares moderated-regression: Some reasons and a remedy. *Psychological Bulletin*, 99, 282–288.
- Mount, M. K., Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the Big Five personality factors. *Journal of Applied Psychology*, 79, 272–280.
- O’Boyle, E. H., Humphrey, R. H., Pollack, J. M., Hawver, T. H., & Story, P. A. (2010). The relation between emotional intelligence and job performance: A meta-analysis. *Journal of Organizational Behavior*, 32, 788–818.
- Oh, I., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *Journal of Applied Psychology*, 96, 762–773.
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology*, 81, 660–679.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88, 879–903.
- Posthuma, R. A., Morgeson, F. P., & Campion, M. A. (2002). Beyond employment interview validity: A comprehensive narrative review of recent research and trends over time. *Personnel Psychology*, 55, 1–81.
- Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe–Crowne social desirability scale. *Journal of Clinical Psychology*, 38, 119–125.
- Roberts, B. W. (2009). Back to the future: Personality and assessment and personality development. *Journal of Research in Personality*, 43, 137–145.
- Robertson, I. T., & Smith, M. (2001). Personnel selection. *Journal of Occupational and Organizational Psychology*, 74, 441–472.
- Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied Psychology*, 83, 634–644.
- Roth, P. L., & Huffcutt, A. I. (2013). A meta-analysis of interviews and cognitive ability: Back to the future? *Journal of Personnel Psychology*, 12, 157–169.
- Roth, P. L., Van Iddekinge, C. H., Huffcutt, A. I., Edison, C. E., Jr, & Schmit, M. J. (2005). Personality saturation in structured interviews. *International Journal of Selection and Assessment*, 13, 261–273.
- Salgado, J. F., & Moscoso, S. (2002). Comprehensive meta-analysis of the construct validity of the employment interview. *European Journal of Work and Organizational Psychology*, 11, 299–326.
- Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, cognition, and personality*, 9, 185–211.
- Schmidt, N., & Ostroff, C. (1986). Operationalizing the “behavioral consistency” approach: Selection test development based on a content-oriented strategy. *Personnel Psychology*, 39, 91–108.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420–428.
- Srull, T. K., & Wyer, R. S. (1989). Person memory and judgment. *Psychological Review*, 96, 58–83.
- Sue-Chan, C., & Latham, G. P. (2005). The situational interview as a predictor of academic and team performance: A study of the

- mediating effects of cognitive ability and emotional intelligence. *International Journal of Selection and Assessment*, *12*, 312–320.
- Townsend, R. J., Bacigalupi, S. C., & Blackman, M. C. (2007). The accuracy of lay integrity assessments in simulated employment interviews. *Journal of Research in Personality*, *41*, 540–557.
- Van Iddekinge, C. H., Raymark, P. H., & Roth, P. L. (2005). Assessing personality with a structured employment interview: Construct-related validity and susceptibility to response inflation. *Journal of Applied Psychology*, *90*, 536–552.
- Wonderlic & Associates. (1992). *Wonderlic personnel test and scholastic level exam: User's Manual*. Libertyville, IL: Wonderlic Personnel Test, Inc.
- Wong, C. S., & Law, K. S. (2002). The effect of leader and follower emotional intelligence on performance and attitude: An exploratory study. *Leadership Quarterly*, *13*, 243–274.

Copyright of Journal of Business & Psychology is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.