Human Performance, 25:335–353, 2012

ISSN: 0895-9285 print/1532-7043 online
DOI: 10.1080/08959285.2012.703732

Routledge
Taylor & Francis Group

Implicit Trait Policies in Multimedia Situational Judgment Tests for Leadership Skills: Can They Predict Leadership Behavior?

Janneke K. Oostrom and Marise Ph. Born
Erasmus University Rotterdam

Alec W. Serlie
GITP/Erasmus University Rotterdam

Henk T. van der Molen
Erasmus University Rotterdam

To explain why situational judgment tests are often correlated with personality measures, Motowidlo, Hooper, and Jackson (2006a, 2006b) developed the implicit trait policy theory. Implicit trait policies are beliefs about causal relationships between personality traits and behavioral effectiveness. Among 180 employees, this field study examined whether a multimedia situational judgment test that was intended to assess leadership skills can capture individual differences in such policies. Furthermore, it was examined whether these implicit trait policies were able to predict leadership behavior. Results confirmed that the situational judgment test was able to capture individual differences in implicit trait policies for Extraversion and Conscientiousness. Furthermore, results showed that implicit trait policies for Extraversion can predict leadership behavior over and above leadership experience and the associated personality trait.

Situational judgment tests (SJTs) are a frequently used selection tool, both in the United States and Europe (McDaniel & Nguyen, 2001; Salgado & Lado, 2000). SJTs typically present job-related situations followed by a number of alternative response options. Applicants are then asked to evaluate the effectiveness of each response option or indicate the likelihood that they would respond in that way (Whetzel & McDaniel, 2009). Meta-analyses have demonstrated that SJTs have useful levels of validity as predictors of job performance (McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001; McDaniel, Hartman, Whetzel, & Grubb, 2007) and that SJTs show substantial correlations with cognitive ability (McDaniel et al., 2001) and with Big Five personality dimensions (McDaniel & Nguyen, 2001; McDaniel et al., 2007). However, there is a paucity of theory regarding the predictive and construct validity of SJTs (Whetzel & McDaniel, 2009).

Correspondence should be sent to Janneke K. Oostrom, Erasmus University Rotterdam, Faculty of Social Sciences, Institute of Psychology, T12-43, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. E-mail: oostrom@fsw.eur.nl

To explain why SJTs are often correlated with measures of personality traits, Motowidlo, Hooper, and Jackson (2006a, 2006b) developed the implicit trait policy (ITP) theory. The ITP theory starts from the assumption that SJTs predict job performance because they measure procedural knowledge, which include a component of general domain knowledge about the costs and benefits of expressing particular personality traits in job-related situations. ITPs are the implicit beliefs of individuals about the effectiveness of different levels of trait expression. For instance, an individual may believe that the expression of Conscientiousness is generally very effective. ITPs are measured by correlating applicants' effectiveness ratings of SJT response options with the level of trait expression of these response options. The central proposition of the ITP theory is that individual differences in personality traits affect judgments of the effectiveness of SJT response options that express those personality traits. Motowidlo et al. (2006a, 2006b) indeed found empirical support for the ITP theory, as they were able to demonstrate that ITPs for Agreeableness, Conscientiousness, and Extraversion are related to individual differences in these personality traits. More recently, Motowidlo and Beier (2010) demonstrated that ITPs are able to predict a performance composite based on supervisor ratings. To measure ITPs, Motowidlo and Beier used an SJT specifically designed for management and administrative positions in telecommunications industry.

Expanding on the study of Motowidlo and Beier (2010), the present study aims to shed light on the predictive validity of ITPs. However, in contrast with Motowidlo and Beier's study and other studies on the predictive validity of ITPs, the present study uses a multimedia SJT that is saturated with a particular construct domain. Saturation refers to the extent to which a particular construct affects a multidimensional measure, like SJTs (Christian, Edwards, & Bradley, 2010; Roth, Bobko, McFarland, & Buster, 2008). For example, when all situations in an SJT describe interpersonal interactions, this SJT is saturated with the construct domain of interpersonal skills. A construct-based approach has the following advantages: (a) the specification of the construct domain helps to reduce contamination due to the measurement of unintended, non-job-relevant constructs; (b) it provides insight into why the SJT is related to the criterion of interest, which increases the ability to generalize findings across time and jobs; and (c) it provides the opportunity to conceptually match the predictor and criterion domain (Christian et al., 2010; Lievens, 2006).

Specifically, in the present study it is examined whether a multimedia SJT saturated with leadership skills is able to capture ITPs for targeted traits and whether these ITPs are able to predict leadership behaviors. The structure of the introduction section is as follows: First, ITP theory is discussed in more detail, followed by an overview of previous research on ITPs. Then, several hypotheses about the relationships between ITPs, personality traits, leadership experience, and leadership behavior are proposed.

ITP THEORY

According to Mowidlo et al. (2006b), the ITP theory has three main assumptions, as depicted in Figure 1. The first assumption is that procedural knowledge as measured by SJTs has a causal effect on job performance. Several meta-analyses have demonstrated the criterion-related validity of SJTs (Christian et al., 2010; McDaniel et al., 2007; McDaniel et al., 2001), lending support for this assumption.

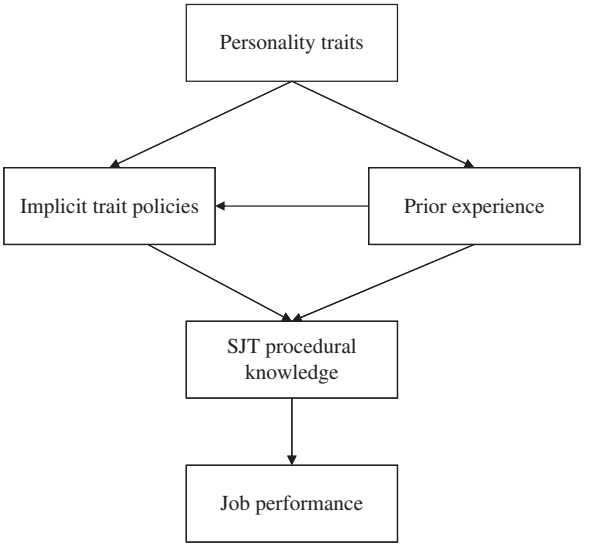

FIGURE 1 Hypothesized model of relationships between personality traits, experience, implicit trait policies, situational judgment test (SJT) performance, and job performance by Motowidlo et al. (2006b).

The second assumption is about the causal effects of personality traits on SJT performance. ITP theory is embedded in social cognition research, which has shown that the judgment of trait-related behaviors of others is determined by the characteristics of the judge him- or herself (e.g., Heider, 1958; Lambert & Wedell, 1991; Markus, Smith, & Moreland, 1985). ITP theory assumes that there are stable differences in individuals' implicit beliefs about the effectiveness of different levels of trait expression (Motowidlo et al., 2006a). An important determinant of how strong one weighs the expression of a particular trait is one's own standing on the trait (Motowidlo et al., 2006b). The reason for this is that individuals tend to believe that their own preferred way to handle a situation is the most effective way. Thus, individual differences in personality traits should affect their judgments of the effectiveness of SJT response options that express those personality traits (Motowidlo et al., 2006b). For example, agreeable individuals judge very agreeable response options in an SJT as more effective than disagreeable individuals. Their ITPs for Agreeableness would, therefore, be represented by a relatively strong positive correlation between their effectiveness ratings of the response options, on one hand, and the degree to which the response options express Agreeableness, on the other hand. This may explain why SJT scores are often correlated with measures of personality traits (McDaniel & Nguyen, 2001; McDaniel et al., 2007).

The third assumption of Motowidlo et al. (2006b) is that ITPs are affected by prior experience. Motowidlo et al. (2006a) argued that SJT performance includes both a component of specific knowledge about effective behavior in a particular job and a component of general knowledge about costs and benefits of expressing particular personality traits in job-related situations. SJTs scores indeed have been found to be related to prior job experience (e.g., Clevenger, Pereira, Wiechmann, Schmitt, & Harvey, 2001; McDaniel & Nguyen, 2001; Weekley & Jones, 1999). People can acquire specific job knowledge only through exposure to a particular job. Yet, general

knowledge about trait effectiveness (ITPs) can be acquired by experience in relevant situations both inside and outside the job. Through these experiences, individuals learn that the expression of certain personality traits is generally more effective than the expression of other personality traits (Motowidlo & Beier, 2010). Individuals will develop ITPs accordingly, independently of their own standing on those traits. It has been suggested that SJTs can only be used for selecting applicants with considerable domain-specific knowledge and experience (e.g., Weekley, Ployhart, & Harold, 2004). As ITPs tap general knowledge acquired independently of specific job experience, they can be used even with applicants who have no relevant job experience.

PREVIOUS RESEARCH ON ITPS

Motowidlo et al. (2006a, 2006b) conducted a number of studies among undergraduates that tested relationships between ITPs as measured with paper-and-pencil SJTs and associated personality traits as measured with the NEO-Five Factor Inventory. Significant correlations were found between ITPs for Agreeableness, Conscientiousness, and Extraversion and the associated personality traits. In a study among 99 undergraduates, Motowidlo et al. (2006b) tested the hypothesis that ITPs for Agreeableness and Extraversion as measured with a paper-and-pencil SJT could predict behavioral expressions of these traits in a role-play exercise. The results partly supported this hypothesis, as they found a significant correlation between ITPs for Agreeableness and role-play Agreeableness scores, even when they partialed out the Agreeableness scores on the NEO-Five Factor Inventory. However, ITPs for Extraversion did not predict behavioral expressions of Extraversion. The explanation of Motowidlo et al. for this latter finding was that the specific facets of Extraversion represented in the ITP measure, for example, taking charge in social situations and standing up for one's own interest, might be different from the specific facets of Extraversion expressed in the role-play exercise, for example, being enthusiastic, unreserved, and talkative.

Recently, in a study among 115 employees, Motowidlo and Beier (2010) demonstrated that ITPs for Agreeableness and Conscientiousness were significantly related to a performance composite score build up by supervisor ratings of 10 different workplace behaviors. To measure these ITPs they used a paper-and-pencil SJT that was designed to predict a number of competencies in managerial and administrative jobs in the telecommunications industry (e.g., leadership, flexibility, sensitivity, communication). In addition, the study of Motowidlo and Beier (2010) showed that scoring keys based on personality expressions as rated by graduate students with no domain-specific job knowledge can produce scores that are related to workplace behaviors. Thus, in contrast to SJT scoring keys that are used to measure procedural knowledge, scoring keys for ITPs do not require experts with considerable domain-specific knowledge and experience.

Among 71 undergraduates, Miller, Smith-Jentsch, and Afek (2008) examined the relationships between ITPs for Agreeableness and Conscientiousness, the personality scale scores of Agreeableness and Conscientiousness, job experience, and peer ratings of typical Agreeableness and Conscientiousness. To measure ITPs, Miller et al. (2008) used an SJT that was designed for a welfare-to-work program. Significant correlations were found between ITPs and the associated personality scale scores. Furthermore, it was found that ITPs for Agreeableness explained unique variance in peer ratings of Agreeableness over and above the Agreeableness scores on

the personality questionnaire. Miller et al. also found support for the hypothesis that ITPs for Agreeableness are affected by undergraduate's prior customer service experience.

PRESENT STUDY

The previous studies on ITPs have shown that it is possible to use SJTs to assess individual differences in ITPs and that ITPs can predict peer trait ratings, behavioral expressions in a role-play exercise, and a composite measure of job performance (Miller et al., 2008; Motowidlo & Beier, 2010; Motowidlo et al., 2006a, 2006b). In the present study, we sought to both replicate and extend these findings by testing the hypothesized ITP model regarding the relationships between ITPs as measured with a multimedia SJT for leadership skills, the associated personality scale scores, leadership experience, and observed leadership behavior. As far as we know, no studies have directly tested the ITP model or examined the relationship between ITPs as measured with an SJT that is saturated with a particular construct domain (in our case, leadership skills) and job behavior in the relevant domain (in our case, leadership behavior). The present study examines whether ITPs can be captured by a multimedia SJT. In a multimedia SJT the situations and response options are presented through the use of video clips.

First, the three assumptions of the ITP theory are tested. The first assumption of Motowidlo et al. (2006b) is that SJT performance is significantly related to job performance. According to Motowidlo et al., traditionally scored SJTs measure procedural knowledge, that is, knowledge about how to behave effectively in situations like those depicted in an SJT. Campbell (1990) emphasized the need to align predictor and criterion domains by using the same underlying construct, as this will enhance the validity of the predictor. To clearly examine the predictive validity of the leadership SJT, one should therefore use a criterion that measures participants' leadership behavior. In the present study, peer ratings and supervisor ratings on observed workplace behaviors regarding the competency of leading and deciding were used. As several meta-analyses already have demonstrated, the criterion-related validity of SJTs (Christian et al., 2010; McDaniel et al., 2007; McDaniel et al., 2007), our first hypothesis is as follows:

H1: Scores on a multimedia SJT for leadership skills will be positively related to peer and supervisor ratings of leading and deciding.

Second, it is examined whether a multimedia SJT for leadership skills is able to capture individual differences in ITPs by examining the relationship between ITPs and the associated personality scale scores. We believe ITPs to be an important determinant of participants' effectiveness ratings of SJT response options representing leadership behaviors, as leadership research has emphasized the role of implicit theories in leadership perceptions (e.g., Epitropaki & Martin, 2004; Lord, De Vader, & Alliger, 1986; Lord, Foti, & De Vader, 1984). These implicit leadership theories represent cognitive schemas of traits and behaviors that followers expect from leaders, such as being enthusiastic and supportive (Epitropaki & Martin, 2004), and have been found to be affected by an individual's own personality traits (e.g., Keller, 1999; Lord et al., 1986). SJT response options were scored on the two personality constructs that are most relevant for leadership behaviors, namely, Extraversion and Conscientiousness (Judge, Bono, Ilies, & Gerhardt, 2002). The second hypothesis therefore is as follows:

H2: ITPs for Extraversion (H2a) and ITPs for Conscientiousness (H2b) as measured with a multimedia SJT for leadership skills will be positively related to personality scale scores of Extraversion and Conscientiousness, respectively.

The third assumption of ITP theory is that experience has a causal effect on individuals' ITPs (Motowidlo et al., 2006a). Through experience, individuals learn that certain personality traits are generally more effective than other personality traits, regardless of their own standing on those traits. Miller et al. (2008) demonstrated that ITPs for Agreeableness are indeed affected by prior job experience ($r = .21$). As the multimedia SJT in the current study measures knowledge about trait effectiveness in leadership behaviors, leadership experience is expected to influence participants' ITPs. Therefore, the third hypothesis is as follows:

H3: Leadership experience will be positively related to ITPs for Extraversion (H3a) and ITPs for Conscientiousness (H3b) as measured with a multimedia SJT for leadership skills.

Furthermore, the predictive validity of ITPs for Extraversion and Conscientiousness as measured with a multimedia SJT of leadership skills is examined. Both personality traits, including Extraversion and Conscientiousness, and leadership experience have been found to be positively related to leadership behavior (e.g., Judge et al., 2002; Thomas & Cheese, 2005). For example, Judge et al. (2002) have meta-analytically demonstrated that leadership is significantly related to Extraversion ($r = .22$) and Conscientiousness ($r = .20$). Therefore, it is interesting to investigate whether the relationship between ITPs and observed leadership behavior can be solely attributed to the causal effects of personality and leadership experience on ITPs, or whether ITPs explain unique variance in observed leadership behavior beyond the variance explained by personality traits and leadership experience. To our knowledge, this study is the first to examine whether ITPs have incremental validity over and above personality and prior job experience in predicting job performance ratings. Based on the findings of Motowidlo et al. (2006a) and Miller et al. (2008) that ITPs for Agreeableness explain a significant part of variance in role-play Agreeableness and peer ratings of Agreeableness beyond the variance explained by explicitly measured Agreeableness, the following can be hypothesized:

H4: ITPs for Extraversion (H4a) and ITPs for Conscientiousness (H4b) as measured with a multimedia SJT for leadership skills will explain a significant part of variance in observed leadership behavior beyond the variance explained by personality scale scores of Extraversion and Conscientiousness and leadership experience.

METHOD

Participants and Procedure

This study was conducted among assessment candidates of a large human resource development (HRD)-consultancy firm in the Netherlands. With the invitation for an assessment, an information brochure and an invitation to participate in the study were sent to all assessment candidates of the HRD-consultancy firm in 2008 and 2009. Subsequently, 450 candidates registered themselves voluntarily to participate. Next, they received an e-mail invitation to complete a multimedia SJT and a job performance scale. The response rate was 40.0% (180 participants). About half of

the participants were employees who were assessed for a promotion opportunity (51.1%), and the other applicants were assessed for a job opportunity (48.9%). The age of the participants varied between 22 and 57 ($M = 38.80$, $SD = 8.44$). One hundred ten participants were male (61.1%) and 70 participants were female (38.9%). Educational levels ranged from high school to master's degree. A large percentage of participants worked in commercial (34.7%) or social (9.7%) sectors. To check whether participation was in any way selective, we compared the age, gender, educational level, and assessment outcome of participants to all candidates of 2008 and 2009 ($N = 13{,}701$). None of these comparisons yielded significant differences.

Measures

Personality Questionnaire

As part of their assessment program at the HRD-consultancy firm, participants completed a 224-item personality questionnaire (Koch, 1998), based on the Five Factor Model of personality (Goldberg, 1990). Participants had to provide their answers to the items on a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). All participants completed the questionnaire within 25 min. We used only participants' scores on the Extraversion (27 items) and the Conscientiousness (40 items) scale. An example of an item of the Extraversion scale is as follows: "Rate yourself on the following statement: Enjoys meeting new people." In a study among 261 candidates (GITP, 2010), the Extraversion and the Conscientiousness scale of the personality questionnaire were found to correlate with the Extraversion ($r = .70$) and the Conscientiousness ($r = .62$) scale of the revised NEO-Personality Inventory (Costa & McCrae, 1992). Coefficient alphas were substantial: $\alpha = .92$ for Extraversion and $\alpha = .92$ for Conscientiousness.

Multimedia SJT

The multimedia SJT was designed to predict leadership skills. The SJT consisted of 17 short, videotaped vignettes of interpersonal situations that leaders are likely to encounter on the job. The vignettes were introduced by a narrator. Each situation froze at an important point, and four possible ways for a leader to handle the situation were presented (for an example of an item, see the appendix). Participants were asked to judge the four response options on a 5-point scale ranging from—(*very ineffective*) to ++ (*very effective*). Participants were instructed to complete the multimedia SJT in a calm environment at home or at work. All participants completed the multimedia SJT within 45 min.

To check whether the multimedia SJT indeed measured procedural knowledge in leadership situations, four experts on situational judgment testing (two female, two male) independently categorized the vignettes of the multimedia SJT. The experts worked in the field of personnel selection and had experience with constructing SJTs. The categorization of the vignettes was based on the classification of leadership behaviors provided by O*Net, which consisted of five categories: (a) making decisions and solving problems; (b) coordinating the work and activities of others; (c) guiding, directing, and motivating others; (d) developing and building teams; and (e) resolving conflicts and negotiating with others. The experts were asked to indicate to which category each vignette belonged by writing down the category number (1 to 5) after each vignette. When they believed that a vignette did not belong to one of the categories of leadership behaviors

they were instructed to write down a 6 after the vignette. One expert indicated that two of the vignettes did not belong to any of the categories of leadership behaviors, the other three experts indicated that each vignette belonged to one of the categories of leadership behaviors. The one-way random effects intraclass correlation (ICC) for absolute agreement was .86, indicating that there was substantial agreement in the categorization of the vignettes among the experts. The experts indicated that two vignettes belonged to the category of making decisions and solving problems; two vignettes to coordinating the work and activities of others; nine vignettes to guiding, directing, and motivating others; two vignettes to developing and building teams; and two vignettes belonged to the category of resolving problems and negotiating with others.

To compute participants' SJT performance, an expert-based scoring method was used (Bergman, Drasgow, Donovan, Henning, & Juraska, 2006). Fifteen experts individually watched the videotaped vignettes and rated the four response alternatives on the same 5-point scale as the participants. The absolute distance between the mean effectiveness ratings of the experts and the participants' effectiveness ratings was calculated for each response alternative. The absolute distances of all responses were summed and extracted from 100, so participants received a higher score if they tended to agree with the experts. Coefficient alpha of the absolute distance scores equaled .88.

To compute participants' ITP scores, four subject matter experts (two male, two female) who worked as the HRD-consultancy firm rated the SJT response options according to the level of Extraversion and Conscientiousness that each expressed. They used a 7-point scale ranging from 1 (*very introverted/very unconscientious*) to 7 (*very extraverted/very conscientious*). Mean trait ratings were computed for each response option. There was substantial agreement between the experts about the level of Extraversion and Conscientiousness the response options expressed (ICC = .76 for Extraversion and ICC = .91 for Conscientiousness). Participants' ITPs for Extraversion and Conscientiousness were represented by the correlation coefficients between the participants' effectiveness ratings of the 68 response options in the SJT and the experts' mean trait ratings. Fisher's z-transformation was used to normalize the correlation coefficients.

Leadership Experience

This variable was measured with the following item: "How many years of leadership experience do you have?" Participants indicated their experience on a 5-point scale ranging from 1 (*no experience*) to 5 (*more than 10 years*).

Criterion Measure

Participants were instructed to ask at least one individual in their direct work environment (preferably their supervisor) to evaluate their job behaviors related to the competency of leading and deciding by filling out a 10-item questionnaire. The competency leading and deciding includes behaviors such as taking control and exercising leadership, initiating actions, giving directions, and taking responsibilities. An example of an item is, "Knows how to motivate employees to achieve their goals." All items had to be rated on a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Performance ratings were obtained for 120 participants. For 49 of these participants only peer ratings were obtained, for 23 of these participants

only supervisor ratings were obtained, and for 48 of these participants both peer ratings and supervisor ratings were obtained. For a small number of participants (24.2%) multiple peer ratings were available. For these participants one peer rating was randomly extracted. This resulted in 97 peer ratings and 71 supervisor ratings. Coefficient alphas were substantial, namely, .87 for the peer ratings and .83 for the supervisor ratings.

RESULTS

Preliminary Analyses

Means, standard deviations, scale reliabilities and correlations between all variables are presented in Table 1. Before testing the hypotheses, we first looked at significant correlations between demographic characteristics and leadership experience, scores on the personality scales, ITPs, and the criterion measure. Gender was significantly related to leadership experience ($r = -.21$, $p < .01$). Male participants ($M = 3.12$, $SD = 1.45$) had more leadership experience than female participants ($M = 2.52$, $SD = 1.31$, $t = 2.75$, $p < .01$). Age was significantly related to leadership experience ($r = .42$, $p < .01$) and to peer ratings of leading and deciding ($r = .23$, $p < .05$). Because of these significant correlations, gender and age were controlled for in the regression analyses. As reported in previous studies (e.g., Harris & Schaubroeck, 1988; Mount, 1984), peer ratings and supervisor ratings were modestly related ($r = .49$, $p < .01$) and mean peer ratings were significantly higher than mean supervisor ratings ($t = 2.00$, $p < .05$). As one of the most consistent findings in the literature on performance appraisal is that the ratings obtained from different sources do not converge (e.g., Harris & Schaubroeck, 1988; Mount, Judge, Scullen, Sytsma, & Hezlett, 1998), we decided to test the ITP model and the hypotheses separately for peer ratings and supervisor ratings of leading and deciding.

To check whether peer ratings and supervisor ratings were obtained from a selective sample, we compared participants who provided performance ratings to participants who did not provide performance ratings on all study variables. None of the comparisons yielded significant differences. Hence, we considered the missing peer ratings and supervisor ratings as missing at random.

Hypotheses Testing

The three assumptions of ITP theory were tested separately for the peer ratings and the supervisor ratings of leading and deciding using structural equation modeling (AMOS 16.0, Arbuckle, 2007). The missing performance ratings were imputed using the expectation-maximization algorithm (Dempster, Laird, & Rubin, 1977). We used several indices to judge the fit of the model to our data, including the chi-square test. Although the chi-square test is the most widely used measure of model fit in organizational research (e.g., Kelloway, 1996), it is also highly sensitive to sample size (Jöreskog & Sörbom, 1993). Hence, we used a number of alternative fit indices, namely, the comparative fit index (CFI), the Tucker–Lewis index (TLI), the standardized root mean square residual (SRMR), and the root mean square error of approximation (RMSEA). CFI and TLI values of .95 or higher, SRMR values of .08 or less, and RMSEA values of .06 or less indicate a relatively good fit between the hypothesized model and the observed data (Hu &

TABLE 1
Means, Standard Deviations, Scale Reliabilities, and Correlations Between Personality Traits, ITPs, and Criterion Measures

	M	*SD*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1. Gender	.39	.49	(—)									
2. Age	38.83	8.44	−.08	(—)								
3. Leadership experience	2.89	1.42	−.21**	.42**	(—)							
4. Personality trait – Extraversion	3.67	0.51	.03	.03	.18*	(.92)						
5. Personality trait – Conscientiousness	3.97	0.46	.10	.01	.12	.25**	(.85)					
6. ITPs Extraversion	−.15	.13	−.13	.13	.23**	.19*	.16	(—)				
7. ITPs Conscientiousness	.46	.12	−.04	.05	.16*	.06	.19*	.24**	(—)			
8. SJT performance	65.32	14.62	.00	.02	.16*	.17*	.20*	.22**	.29**	(.88)		
9. Peer ratings leading and deciding	3.63	0.57	−.07	.23*	.38**	.13	.36**	.31**	.22**	.28**	(.87)	
10. Supervisor ratings leading and deciding	3.41	0.52	−.07	.04	.20	.10	.21	.26*	.26*	.23*	.49**	(.83)

Note. Scale reliabilities are presented on the diagonal. Personality scales and ratings of workplace behaviors were measured on a 5-point scale. Gender (0 = male, 1 = female) and leadership experience (1 = *no experience*, 2 = *less than 1 year*, 3 = *1 to 5 years*, 4 = *6 to 10 years*, and 5 = *more than 10 years*) were coded. Implicit trait policies (ITPs) are presented by a correlation coefficient, and Situational Judgment Test (SJT) performance is measured on a scale with a maximum of 100. $N = 180$ for demographic characteristics, personality scales, and ITPs. $N = 97$ for peer ratings of leading and deciding. $N = 71$ for supervisor ratings of leading and deciding. The correlation between peer ratings and supervisor ratings of leading and deciding is based on $N = 48$.

$^{*}p < .05$. $^{**}p < .01$.

Bentler, 1999), whereas CFI and TLI values of .90 or higher (Marsh, Hau, & Wen, 2004), SRMR values of .10 or less (Hu & Bentler, 1995), and RMSEA values of .08 or less (MacCallum, Browne, & Sugawara, 1996) indicate an acceptable fit.

Because previous studies have found intercorrelations among ITPs for targeted traits (e.g., Motowidlo et al., 2006a) and among personality traits (e.g., Van der Linden, Te Nijenhuis, & Bakker, 2010), these variables were allowed to covary. The hypothesized model showed an acceptable fit for both the peer ratings of leading and deciding, $\chi^2(9) = 18.78$, $p < .05$, CFI = .90, TLI = .86, SRMR = .06, RMSEA = .07, and the supervisor ratings of leading and deciding, $\chi^2(9) = 17.16$, $p < .05$, CFI = .91, TLI = .87, SRMR = .06, RMSEA = .07. Figure 2 presents the model for both performance ratings, with all estimated path coefficients being positive and significant.

Consistent with H1, which stated that scores on a multimedia SJT for leadership skills would be positively related to peer and supervisor ratings of leading and deciding, a positive standardized path coefficient of .25 ($p < .01$) between SJT performance and peer ratings of leading and deciding and a positive standardized path coefficient of .15 ($p < .05$) between SJT performance and supervisor ratings of leading and deciding was found.

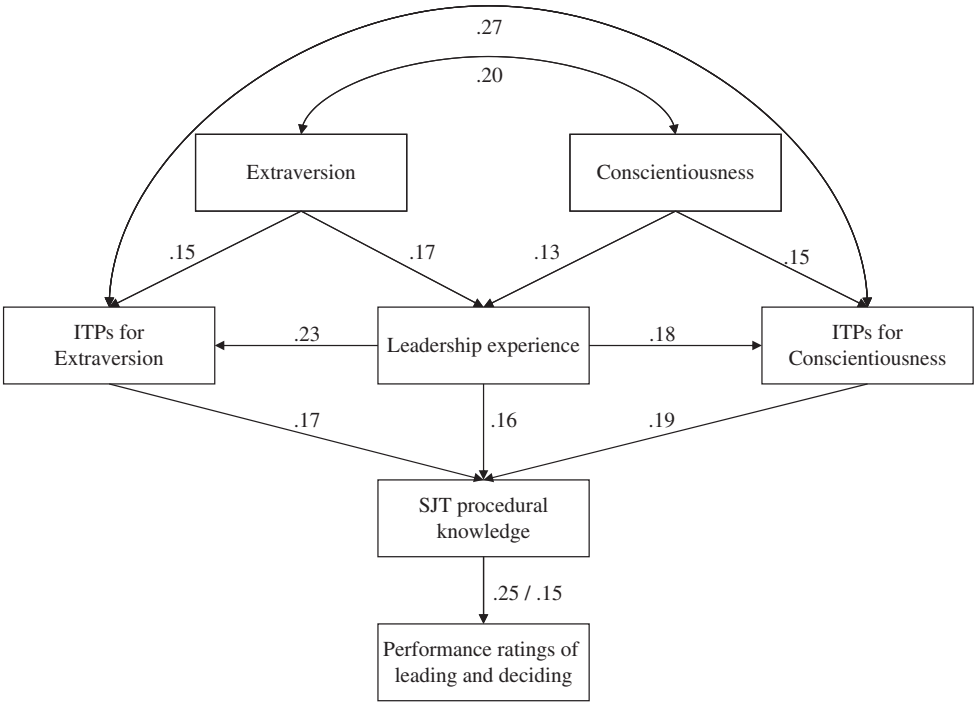

FIGURE 2 Standardized path coefficients for relationships between personality traits, leadership experience, implicit trait policies for Extraversion, situational judgment test (SJT) performance, and observed leadership behavior. *Note.* Regarding the path coefficients between SJT procedural knowledge and performance ratings of leading and deciding, the first coefficient concerns the peer ratings of leading and deciding and the second coefficient concerns the supervisor ratings of leading and deciding. All estimated path coefficients are significant at $p < .05$.

H2, stating that ITPs for Extraversion (H2a) and ITPs for Conscientiousness (H2b) would be positively related to personality scale scores of Extraversion and Conscientiousness, respectively, was supported. A positive standardized path coefficient of .15 ($p < .05$) was found between both the personality trait Extraversion and ITPs for Extraversion and the personality trait Conscientiousness and ITPs for Conscientiousness.

H3, which stated that leadership experience would be positively related to ITPs for Extraversion (H3a) and ITPs for Conscientiousness (H3b), was also supported. A positive standardized path coefficient of .23 ($p < .01$) was found between leadership experience and ITPs for Extraversion and a positive standardized path coefficient of .18 ($p < .05$) was found between leadership experience and ITPs for Conscientiousness.

A series of hierarchical regression analyses was conducted to test H4, stating that ITPs for Extraversion (H4a) and ITPs for Conscientiousness (H4b) as measured with a multimedia SJT for leadership skills would explain a significant part of variance in observed leadership behavior beyond the variance explained by personality scale scores of Extraversion and Conscientiousness and leadership experience. Gender and age were entered in the first step, leadership experience and Extraversion or Conscientiousness as measured with the personality questionnaire in the second step, and ITPs for Extraversion or ITPs for Conscientiousness were entered in the final step. Tables 2 and 3 present the results for the ITPs for Extraversion and the ITPs for Conscientiousness, respectively.

ITPs for Extraversion were able to explain additional variance in peer ratings of leading and deciding ($\beta = .18$, $p < .05$, $\Delta R^2 = .03$, $p < .05$) over and above leadership experience and the personality trait Extraversion, but not in supervisor ratings of leading and deciding ($\beta = .14$, *ns*, $\Delta R^2 = .02$, *ns*). Regarding peer ratings, a significant beta weight was also found for leadership experience ($\beta = .21$, $p < .01$). Taken together, the control variables and the predictors explained

TABLE 2
Incremental Validity of ITPs for Extraversion Over and Above Leadership Experience and Extraversion in Predicting Leading and Deciding

	Peer Ratings[a]		*Supervisor Ratings*[b]	
	β	ΔR^2	β	ΔR^2
Step 1				
Gender	.02		−.01	
Age	.05	.03	−.04	.01
Step 2				
Leadership experience	.21**		.10	
Extraversion	.02	.05**	.01	.01
Step 3				
ITPs Extraversion	.18*	.03*	.14	.02
R^2	.11		.04	
F	4.28**		1.24	

Note. Standardized regression weights are for final step. Gender (0 = male, 1 = female) and leadership experience (0 = *no experience*, 1 = *less than 1 year*, 2 = *1 to 5 years*, 3 = *6 to 10 years*, and 4 = *more than 10 years*) were coded. ITP = implicit trait policy.

[a] $N = 97$. [b] $N = 71$.

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

TABLE 3
Incremental Validity of ITPs for Conscientiousness Over and Above Leadership Experience and Conscientiousness in Predicting Leading and Deciding

	Peer Ratings[a]		*Supervisor Ratings*[b]	
	β	ΔR^2	β	ΔR^2
Step 1				
Gender	−.01		−.02	
Age	.07	.03	−.03	.01
Step 2				
Leadership experience	.22**		.10	
Conscientiousness	.12	.07**	.06	.02
Step 3				
ITPs Conscientiousness	.10	.01	.14	.02
R^2	.11		.04	
F	4.16**		1.51	

Note. Standardized regression weights are for final step. Gender (0 = male, 1 = female) and leadership experience (0 = no experience, 1 = less than one year, 2 = one to five years, 3 = 6–10 years, and 4 = more than 10 years) were coded. R^2 and ΔR^2 may appear inconsistent due to rounding. ITP = implicit trait policy.

[a]$N = 97$. [b]$N = 71$.

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

11% of the variance ($F = 4.28$, $p < .01$) in peer ratings of leading and deciding and 4% of the variance ($F = 1.24$, *ns*) in supervisor ratings of leading and deciding. These results partly support H4a.

ITPs for Conscientiousness were not able to explain additional variance in peer ratings of leading and deciding ($\beta = .10$, *ns*, $\Delta R^2 = .01$, *ns*) and supervisor ratings of leading and deciding ($\beta = .14$, *ns*, $\Delta R^2 = .02$, *ns*) over and above leadership experience and the personality trait Conscientiousness. Regarding peer ratings, a significant beta weight was found for leadership experience ($\beta = .22$, $p < .01$). The control variables and the predictors explained a significant part of the variance in peer ratings of leading and deciding ($R^2 = .11$, $F = 4.16$, $p < .01$). Based on these findings, H4b could not be supported.

DISCUSSION

The aim of the present study was to examine the relationships between ITPs as measured with a multimedia SJT for leadership skills, personality scale scores, leadership experience, and leadership behavior. Furthermore, the present study examined whether ITPs have incremental validity over and above personality and prior job experience in predicting job behavior in the relevant domain. Results confirmed that a multimedia SJT for leadership skills can be used to measure individual differences in ITPs and that those ITPs are able to predict leadership behavior over and above leadership experience and personality traits. Each of the findings is discussed next.

The first assumption of ITP theory is that procedural knowledge as measured by SJTs has a causal effect on job performance (Motowidlo et al., 2006b). In line with several meta-analyses

that have demonstrated the criterion-related validity of SJTs (Christian et al., 2010; McDaniel et al., 2007; McDaniel et al., 2001), the results of the present study demonstrated that performance on the multimedia SJT for leadership is able to predict observed leadership behavior. The first hypothesis, which stated that scores on a multimedia SJT for leadership skills would be positively related to peer and supervisor ratings of leading and deciding, therefore was supported.

The second assumption of ITP theory is that individual differences in implicit beliefs about the effectiveness of different levels of trait expression is affected by one's own standing on the trait (Motowidlo et al., 2006b). The hypothesis with respect to this assumption, which stated that ITPs for Extraversion and Conscientiousness would be positively related to the personality scale scores of Extraversion and Conscientiousness, respectively, was supported. Participants' level of Extraversion and Conscientiousness as measured with a personality scale was found to be related to their ITPs for these traits. The present results are in line with Motowidlo et al. (2006a, 2006b) and Miller et al. (2008), who showed that ITPs for Agreeableness, Conscientiousness, and Extraversion are related to the associated personality traits.

Support was also found for the third hypothesis that ITPs are affected by job experience. Our results demonstrated that leadership experience explained a significant part of variance in ITPs for Extraversion and ITPs for Conscientiousness. In other words, employees who had more experience as a leader held stronger beliefs about the effectiveness of Extraversion and Conscientiousness in the leadership behaviors that were demonstrated in the SJT. Miller et al. (2008) already demonstrated that prior customer service experience was related to ITPs for Agreeableness. These findings demonstrate that individuals develop implicit beliefs about effective ways to behave on the job by prior experience.

In sum, the results demonstrated that a multimedia SJT for leadership skills indeed can be used to capture ITPs for targeted personality traits. The present study is the first study that has found support for Motowidlo et al's (2006a, 2006b) ITP theory by assessing the complete ITP model through structural equation analyses. The findings provide support for the idea that SJTs predict job performance because they measure procedural knowledge, which include a component of general domain knowledge about the costs and benefits of expressing particular personality traits in job-related situations. The ITP theory provides practitioners an alternative scoring procedure for SJTs, by which this general domain knowledge about the costs and benefits of expressing particular personality traits can be measured. Yet, for this alternative scoring procedure to have practical value, the scores produced by this procedure (ITPs) should explain unique variance in job performance. For this reason, the present study also examined the predictive and incremental validity of ITPs over and above leadership experience, and personality traits.

The fourth hypothesis, which stated that ITPs for Extraversion and ITPs for Conscientiousness would explain a significant part of variance in observed leadership behavior beyond the variance explained by personality scale scores of Extraversion and Conscientiousness and leadership experience, was partly supported. Results demonstrated that ITPs for Extraversion explained 3% of the variance in peer ratings of leadership behavior over and above leadership experience and the personality scale score. Thus, the relationship between ITPs for Extraversion and leadership behavior is not attributable solely to the causal effects of the personality trait of Extraversion and leadership experience on ITPs.

Results further showed that ITPs for Extraversion explained 2% of the variance in supervisor ratings of leadership behavior and ITPs for Conscientiousness explained 1% of the variance in peer ratings and 2% of the variance in supervisor ratings of leadership behavior over and

above leadership experience and the personality scale scores. However, these results were not statistically significant. Note that the difference of these explained variances and the explained variance in peer ratings of leadership behavior by ITPs for Extraversion is fairly minor and therefore may be of little practical importance. Motowidlo et al. (2006a) and Miller et al. (2008) did find that ITPs are able to incrementally predict role-play behavior and peer ratings over and above personality scale scores. The small sample of peers and supervisors in our study who provided rating of participants' leadership behavior led to low power. This could explain the findings of the present study that ITPs for Extraversion did not incrementally predict supervisor ratings and that ITPs for Conscientiousness did not incrementally predict peer and supervisor ratings of leadership behavior.

Miller et al. (2008) found that, as opposed to ITPs for Agreeableness, ITPs for Conscientiousness did not explain unique variance in peer ratings of this trait over and above the personality scale score. As SJTs most often consist of vignettes that show interpersonal situations, it is possible that they are better able to capture ITPs in traits that are most strongly related to interpersonal interactions, namely, Extraversion and Agreeableness (McCrae & Costa, 1989). For this reason, ITPs for Extraversion and ITPs for Agreeableness might be more predictive of behavioral expressions of these traits. This could also explain why significant incremental validity for ITPs for Extraversion was found but not for ITPs for Conscientiousness. Another explanation could be that the personality scale scores of Conscientiousness had substantial validity for peer ratings and supervisor ratings of leading and deciding ($r = .36$ and $r = .21$, respectively), whereas the personality scale scores of Extraversion had only low validity peer ratings and supervisor ratings of leading and deciding ($r = .13$ and $r = .10$, respectively). So, there was more variance in performance ratings left to explain by ITPs over and above the personality scale scores of Extraversion than over and above the personality scale scores of Conscientiousness. Future studies should examine whether the results are due to the type of SJT, in terms of the construct it intends to measure, or to the small sample size. Therefore, it is important to replicate the present findings in a larger sample using SJTs that measure various constructs.

The present study confirms that ITPs as measured with a multimedia SJT for leadership can be predictive of peer ratings of leadership behavior over and above leadership experience and personality scale scores. ITPs, therefore, seem a useful predictor of job performance. There are several advantages of using SJTs to measure ITPs instead of procedural knowledge. First, in contrast to SJT scoring keys that are used to measure procedural knowledge, scoring keys for ITPs do not require experts with considerable domain-specific knowledge and experience. Second, as ITPs tap general knowledge acquired independently of specific job experience, they can be used even with applicants who have no relevant job experience. Third, as SJT performance taps both specific job knowledge and general knowledge, whereas ITPs tap only general knowledge, the validity of ITPs for targeted traits may be more generalizable across job domains than the validity of traditionally scored SJTs.

Limitations and Suggestions for Future Research

There are a number of limitations to the study that should be noted. The first limitation involves the subjective performance rating. Although assessments of workplace behaviors most commonly consist of ratings made by participants' supervisors, peers, or subordinates, these ratings are

potentially biased by selective recall or halo effects. Hogan, Curphy, and Hogan (1994) therefore argued that leadership behavior should be objectively measured in terms of team, group, or organizational effectiveness. Despite this obvious limitation, we believe that there are also several strengths concerning the performance ratings used in the present study, such as the inclusion of multiple raters. Nevertheless, we advise future research to include objectively measured performance outcomes.

The second limitation concerns the content of our criterion measure. Although we have made an effort to align the predictor and criterion domain by using the same underlying construct, admittedly the predictor and criterion could have overlapped even more (e.g., by focusing the criterion measure on the O*Net classification used to cluster the SJT vignettes). As predictor-criterion alignment enhances the validity of the predictor, we encourage future studies regarding the validity of ITPs to carefully consider the overlap between the constructs the SJT aims to measure and the criterion. Furthermore, the items of the criterion measure aimed to measure whether participants have adequate knowledge and skills regarding the competency of leading and deciding and not whether participants actually show this knowledge and skills in their daily work. As knowledge and skills are not directly observable, it might have been difficult for peers and supervisors to evaluate participants' competency of leading and deciding. Future studies should therefore use criterion measures with observable behaviors as items.

The third limitation involves the voluntary nature of study participation. A comparison of age, gender, educational level, and assessment outcome of candidates who actually participated in the study to all other candidates yielded no significant differences, but it remains plausible that motivational difference between the participants and other candidates exists. These motivational differences could have affected performance on the predictors but also on the criterion measure. The employees that participated were actively seeking feedback on their skills and competencies. As in most studies in which performance ratings from multiple raters are obtained (e.g., Murphy & Shiarella, 1997; Scullen, Mount, & Goff, 2000), participants in this study selected the peers and supervisors who rated them. Because people may be motivated to select those peers and supervisors who will rate them positively, this might have resulted in some amount of range restriction. Furthermore, we do not know how the raters perceived the feedback-seeking behavior of the participants. These issues appear to be worthwhile topics for future research.

The present study and previous studies on ITPs have shown that ITPs are related to job experience, personality traits, and relevant job behaviors. However, not much is known about how ITPs are related to other constructs. For example, Motowidlo and Beier (2010) proposed that cognitive ability has a direct effect on ITPs, as more intelligent individuals are expected to learn more quickly what trait expressions are more effective in situations like those depicted in an SJT. However, the relationship between cognitive ability and ITPs has not yet been empirically tested. Future studies should examine the relationship between ITPs, cognitive ability, and other established constructs, such as emotional intelligence or social skills.

ACKNOWLEDGMENTS

We thank Paul E.A.M. van der Maesen de Sombreff and Barend P.N. Koch for providing the multimedia SJT materials and Nienke van Bentveld for her help with the data collection. We acknowledge Stephan J. Motowidlo for his valuable suggestions on an earlier version of this article.

REFERENCES

Arbuckle, J. L. (2007). *Amos 16.0*. Chicago: SPSS.

Bergman, M. E., Drasgow, F., Donovan, M. A., Henning, J. B., & Juraska, S. E. (2006). Scoring situational judgment tests: Once you get the data, your troubles begin. *International Journal of Selection and Assessment*, *14*, 223–235.

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 687–732). Palo Alto, CA: Consulting Psychologists Press.

Christian, M. S., Edwards, J. C., & Bradley, J. C. (2010). Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology*, *63*, 83–117.

Clevenger, J., Pereira, G. M., Wiechmann, D., Schmitt, N., & Harvey, V. S. (2001). Incremental validity of situational judgment tests. *Journal of Applied Psychology*, *86*, 410–417.

Costa, P. T., Jr., & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual. Odessa, FL: Psychological Assessment Resources.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, *39*, 1–38.

Epitropaki, O., & Martin, R. (2004). Implicit leadership theories in applied settings: Factor structure, generalizability, and stability over time. *Journal of Applied Psychology*, *89*, 293–310.

GITP. (2010). *Manual G5R personality questionnaire*. Unpublished manuscript.

Goldberg, L. R. (1990). An alternative description of personality: The Big-Five factor structure. *Journal of Personality and Social Psychology*, *59*, 1216–1229.

Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, *41*, 43–62.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Hogan, R., Curphy, G. J., & Hogan, J. (1994). What we know about leadership: Effectiveness and personality. *American Psychologist*, *49*, 493–504.

Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Thousand Oaks, CA: Sage.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Chicago: Scientific Software International.

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, *87*, 765–778.

Keller, T. (1999). Images of the familiar: Individual differences and implicit leadership theories. *The Leadership Quarterly*, *10*, 589–607.

Kelloway, E. K. (1996). Common practices in structural equation modeling. In C. L. Copper, & I. Robertson (Eds.), *International review of industrial and organizational psychology* (pp. 141–180). Chichester, UK: Wiley.

Koch, B. P. N. (1998). *Handleiding G5*. Unpublished manuscript:

Lambert, A. J., & Wedell, D. H. (1991). The self and social judgment: Effects of affective reaction and "own position" on judgments of unambiguous and ambiguous information about others. *Journal of Personality and Social Psychology*, *61*, 884–897.

Lievens, F. (2006). International situational judgment tests. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 279–300). Mahwah, NJ: Erlbaum.

Lord, R. G., De Vader, C. L., & Alliger, G. M. (1986). A meta-analysis of the relation between personality and leadership perceptions: An application of validity generalization procedures. *Journal of Applied Psychology*, *71*, 402–410.

Lord, R. G., Foti, R. J., & De Vader, C. L. (1984). A test of leadership categorization theory: Internal structure, information processing, and leadership perceptions. *Organizational Behavior and Human Performance*, *34*, 343–378.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130–149.

Markus, H., Smith, J., & Moreland, R. L. (1985). Role of the self-concept in the perception of others. *Journal of Personality and Social Psychology*, *49*, 1494–1512.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*, 320–341.

McCrae, R. R., & Costa, P. T. (1989). The structure of interpersonal traits: Wiggins's circumplex and the five-factor model. *Journal of Personality and Social Psychology*, *56*, 586–595.

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

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

McDaniel, M. A., & Nguyen, N. T. (2001). Situational judgment tests: A review of practice and constructs assessed. *International Journal of Selection and Assessment*, *9*, 103–113.

Miller, D., Smith-Jentsch, K. A., & Afek, A. (2008, April). *Situational judgment tests as measures of implicit trait policies*. Paper presented at the 24th Annual Conference of the Society for Industrial and Organizational Psychology, New Orleans, LA.

Motowidlo, S. J., & Beier, M. E. (2010). Differentiating specific job knowledge from implicit trait policies in procedural knowledge measured by a situational judgment test. *Journal of Applied Psychology*, *95*, 321–333.

Motowidlo, S. J., Hooper, A. C., & Jackson, H. L. (2006a). Implicit policies about relations between personality traits and behavioral effectiveness in situational judgment items. *Journal of Applied Psychology*, *91*, 749–761.

Motowidlo, S. J., Hooper, A. C., & Jackson, H. L. (2006b). A theoretical basis for situational judgment tests. In J. A. Weekley, & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 57–81). Mahwah, NJ: Erlbaum.

Mount, M. K. (1984). Psychometric properties of subordinate ratings of managerial performance. *Personnel Psychology*, *37*, 687–702.

Mount, M. K., Judge, T. A., Scullen, S. E., Sytsma, M. R., & Hezlett, S. A. (1998). Trait, rater, and level effects in 360-degree performance ratings. *Personnel Psychology*, *51*, 557–576.

Murphy, K. R., & Shiarella, A. H. (1997). Implications of the multidimensional nature of job performance for the validity of selection tests: Multivariate frameworks for studying test validity. *Personnel Psychology*, *50*, 823–854.

Roth, P., Bobko, P., McFarland, L., & Buster, M. (2008). Work sample tests in personnel selection: A meta-analysis of black and white differences in overall and exercise scores. *Personnel Psychology*, *61*, 637–661.

Salgado, J. F., & Lado, M. (2000, April). *Validity generalization of video tests for predicting job performance ratings*. Paper presented at the 15th Annual Conference of the Society of Industrial and Oganizational Psychology, New Orleans, LA.

Scullen, S. E., Mount, M. K., & Goff, M. (2000). Understanding the latent structure of job performance ratings. *Journal of Applied Psychology*, *85*, 956–970.

Thomas, R. J., & Cheese, P. (2005). Leadership: Experience is the best teacher. *Strategy & Leadership*, *33*, 24–29.

Van der Linden, D., Te Nijenhuis, J., & Bakker, A. B. (2010). The general factor of personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of Research in Personality*, *44*, 315–327.

Weekley, J. A., & Jones, C. (1999). Further studies of situational tests. *Personnel Psychology*, *52*, 679–700.

Weekley, J. A., Ployhart, R. E., & Harold, C. M. (2004). Personality and situational judgment tests across applicant and incumbent settings: An examination of validity, measurement, and subgroup differences. *Human Performance*, *17*, 433–461.

Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review*, *19*, 188–202.

APPENDIX

Example of a Multimedia Situational Judgment Test Item

Description of Situation

"Two coworkers are supposed to work together on a project. However, the collaboration between the two coworkers is not going that well. One of the coworkers is complaining to the supervisor."

Coworker

"Can I speak to you for a moment?"

Supervisor

"Of course."

Coworker

"I can no longer work this way! Peter is impossible to work with! He doesn't consult me on the project, when we have an agreement he doesn't stick to it, and he only does what he thinks is best. It doesn't work that way. I've tried talking to him about this problem, but he does not want to listen to me. I'm sorry, but I refuse to work with him on this project any longer!"

Possible Reactions

1) "Well, I can't just delegate this project to someone else. You can at least try working with him in a professional way. There are many colleagues who don't like each other, but are still capable of working together."
2) "Well, that's impossible! We are all professionals and you cannot quit before finishing the project. I expect you will resolve this problem together. The project needs to be finished, you should understand that."
3) "Oh my . . . how annoying. I understand the problem has escalated and you can no longer work this way. To be honest, the project has to be finished on time. Maybe we can look for a solution. Do you have any idea how this problem can be resolved?"
4) "Too bad, that the collaboration is not going well. . . . I propose that you tell me everything that's bothering you, so we can look for a possible solution for this problem. Is that all right with you?"