The Effects of Response Instructions on Situational Judgment Test Performance and Validity in a High-Stakes Context

Filip Lievens
Ghent University

Paul R. Sackett
University of Minnesota, Twin Cities Campus

Tine Buyse
Ghent University

This study fills a key gap in research on response instructions in situational judgment tests (SJTs). The authors examined whether the assumptions behind the differential effects of knowledge and behavioral tendency SJT response instructions hold in a large-scale high-stakes selection context (i.e., admission to medical college). Candidates (N = 2,184) were randomly assigned to a knowledge or behavioral tendency response instruction SJT, while SJT content was kept constant. Contrary to prior research in low-stakes settings, no meaningfully important differences were found between mean scores for the response instruction sets. Consistent with prior research, the SJT with knowledge instructions correlated more highly with cognitive ability than did the SJT with behavioral tendency instructions. Finally, no difference was found between the criterion-related validity of the SJTs under the two response instruction sets.

Keywords: situational judgment test, response instructions, high-stakes testing

Situational judgment tests (SJTs) have been the focus of considerable attention in recent years. The presence of an edited volume in the SIOP Frontiers series devoted entirely to the topic (Weekley & Ployhart, 2006) signals the degree of interest in and research on the topic. SJTs are of interest for a variety of reasons, including their conceptual appeal as measures of judgment and problem solving in applied settings, their promise as measures offering incremental validity over established measures in the ability and personality domain, and their potential role in reducing adverse impact against protected groups in selection settings. As Landy (2007) noted, “They seem to represent psychometric alchemy (adverse impact is down, validity is up), they seem to assess practically important KSAOs [knowledge, skills, abilities, and other characteristics], and asseesees like them” (p. 418).

Given this interest, researchers on many fronts have attempted to understand features affecting construct-related and criterion-related validity of SJTs. One such feature is the response instructions given to test takers. In SJTs, several types of response instructions are typically used. For instance, Ployhart and Ehrhart (2003) identified six different types of SJT instructions. Recently, McDaniel, Hartman, Whetzel, and Grubb (2007) classified the various response instructions into a more parsimonious response instruction taxonomy. They made a distinction between SJTs with a knowledge format (e.g., “What is the best answer?”) and SJTs with a behavioral tendency format (e.g., “What are you most likely to do?”).

The present study contributes to prior research by testing whether the common assumptions behind the differential effects of knowledge and behavioral tendency SJT response instructions hold in an actual high-stakes selection context. These differences are examined in terms of effects of SJT instructions on mean scores, correlation with cognitive ability, and criterion-related validity.

Study Background

Prior Research on SJT Response Instructions

Effect on mean scores. To elucidate our conceptual understanding about SJT response instructions, McDaniel et al. (2007) framed them into the typical versus maximal performance distinction. Building on Cronbach (1984), their central premise was that SJTs with knowledge response instructions measure maximal performance. Similar to cognitive ability or job knowledge tests, SJTs with knowledge response instructions ask candidates to show whether they know what the most effective answer is. Conversely, it is posited that SJTs with behavioral tendency instructions measure typical performance because they require candidates to report how they typically behave, which is similar to noncognitive inventories.

On the basis of this central premise, McDaniel et al. (2007) posited that the type of response instructions affects the amount of response distortion in SJTs. As measures of typical performance, SJTs with behavioral tendency instructions are assumed to be more
susceptible to unconscious (self-deception) and conscious (impression management) response distortion than SJTs with knowledge instructions, which are considered to be maximal performance measures. There is some empirical evidence for these arguments. Nguyen, Biderman, and McDaniel (2005) administered an SJT with both knowledge and behavioral tendency instructions to a sample of students, who had to complete both versions under both honest” and fake-good conditions. Results showed that SJTs with behavioral tendency instructions were more fakable (average $d = .25$) than SJTs with knowledge instructions (average $d = .06$). Although these findings show that test takers can fake more readily with behavioral tendency instructions than with knowledge instructions, they do not address the question of whether test takers in high-stakes testing actually do respond differently to the two types of instructions.

Another interesting feature of Nguyen et al.’s (2005) study is that under honest conditions, scores in the knowledge instruction condition were about 1 SD higher than scores in the behavioral tendency instruction condition. Thus, there is evidence that, at least for the SJT under examination in that study, there is a considerable difference between what one should do and what one would do.

**Effect on cognitive loading.** The type of response instruction may also affect the cognitive loading of the SJTs. As defined by Whetzel, McDaniel, and Nguyen (2008), cognitive loading of an SJT refers to the extent to which the SJT is correlated with cognitive ability. Because SJTs with knowledge instructions are considered maximal performance measures, we posit that they correlate more with cognitive ability measures. Conversely, as typical performance measures, SJTs with behavioral tendency instructions are assumed to be less correlated with cognitive ability measures. McDaniel et al.’s (2007) meta-analysis confirmed that SJTs with knowledge instructions correlated more highly with cognitive ability tests (.35) than did SJTs with behavioral tendency instructions (.19). Conversely, SJTs with behavioral tendency instructions correlated more highly with agreeableness (.37), conscientiousness (.34), and emotional stability (.35) than did SJTs with knowledge instructions (.19, .24, and .12, respectively).

**Effect on criterion-related validity.** Response instructions may affect the criterion-related validity of SJTs. So far, two competing arguments have been put forward. Some researchers (McDaniel et al., 2007) proposed that knowledge instructions are more valid because the SJT basically measures job knowledge, with the latter being a good predictor of job performance. Other researchers (Ployhart & Ehrhart, 2003) offered the possibility that behavioral tendency instructions may be more valid because of the behavioral consistency explanation (intended behavior predicts future behavior).

The issue of fakability adds another layer to the question of the relative validity of knowledge versus behavioral tendency instructions. If knowledge instructions create a maximum performance setting, then test takers cannot improve their score by faking (Nguyen et al., 2005). One either knows the appropriateness of various responses or one does not. Under knowledge instructions, any differences in test scores between high-stakes applicant settings and low-stakes research settings would not be due to faking. Differences are possible for other reasons, such as failure to devote attentional resources to the testing in the low-stakes setting. Conversely, the potential for differences between high-stakes and low-stakes settings under behavioral tendency instructions is considerable. Under high-stakes conditions, test takers are motivated to score highly; thus, there is a considerable dilemma for the test taker who clearly recognizes a particular course of action as what one should do but has the self-insight to know that he or she would not do this (“I know the best thing to do is to act immediately and not procrastinate, but I know that I am likely to procrastinate anyway”). It is likely that some test takers resolve this dilemma by reporting what they should do, whereas others report what they would do. Thus, it is possible that behavioral tendency instructions have differing effects in high-stakes and low-stakes conditions and that the relative effectiveness of knowledge instructions and behavioral tendency instructions differs in high-stakes and low-stakes settings.

Empirical evidence to date regarding the effects of response instructions on SJT criterion-related validities has been mixed. The meta-analysis of McDaniel et al. (2007) examined the effect of SJT response instructions on the criterion-related validity of SJTs. No significant differences were found. When McDaniel et al. limited the sample of studies in their meta-analysis to studies wherein SJT content was held constant, SJTs with knowledge instructions had higher criterion-related validity than SJTs with behavioral tendency instructions. However, Ployhart and Ehrhart (2003) drew the opposite conclusion. In their experimental study, they compared the criterion-related validity of knowledge versus behavioral tendency instructions on exactly the same five-item SJT of college student success. SJTs with behavioral tendency instructions significantly outperformed SJTs with knowledge instructions for predicting grade point average (GPA), self-rated performance, and peer-rated performance.

**Limitations of Prior Research**

McDaniel et al. (2007) highlighted two key limitations of their meta-analytic work. The first is the inability to keep the SJT content constant, resulting in a lack of method–construct distinction (Arthur & Villado, 2008). Hence, it is possible that studies using SJTs with knowledge instructions differed on some unknown set of features (e.g., the content of the items) from studies using SJTs with behavioral tendency instructions. Granted, McDaniel et al. analyzed SJT–performance correlations separately for three studies in which content was held constant (out of a total of 118 SJT–performance correlations). However, they did conclude that additional research manipulating response instructions with content held constant would be valuable. We achieved this in the current study.

The second key limitation is that almost all studies included in McDaniel et al.’s (2007) meta-analysis used incumbents. Only 4 of the 118 studies did not do so. Clearly, issues of motivation to self-present are quite different in low-stakes incumbent settings than in high-stakes applicant settings. Therefore, it remains to be seen whether the differences between SJT response instructions observed in prior research are still obtained in a high-stakes selection context. This is a key missing piece in the knowledge of SJT response instructions. In the present study, we focus on such a high-stakes applicant setting.

**Present Study**

In this study, we had the unique opportunity to experimentally manipulate the response instructions in a high-stakes context: an
admission exam for medical studies in Belgium. This was possible
because the response instruction issue was a long-standing debate
in the scientific commission overseeing the admission exam. To
settle this issue, the commission agreed to randomly assign stu-
dents to one of two conditions: SJT with knowledge instructions
and SJT with behavioral tendency instructions. Apart from the
response instructions, the SJT content was exactly the same across
the two conditions. A contingency plan was in place to equate
scores should substantial mean score differences be found. This
study’s design permits us to examine the effects of SJT instruc-
tions on mean scores, correlation with cognitive ability, and
criterion-related validity.

Hypotheses/Research Questions

In considering the likely effects of knowledge versus behavioral
tendency instructions, we found it useful to think about two fea-
tures of the situation, which create a $2 \times 2$ set of scenarios when
crossed. Factor 1 deals with faking. Note that with SJTs, faking
means giving a knowledge response, even if it is different from a
behavioral tendency response. Under knowledge instructions, fak-
ing is generally not an issue: One is asked which response is most
or least effective, and one responds accordingly. We should note
that it is conceivable that there is a difference between what the
person views as the best response and what the person believes the
organization views as the best response; therefore, at that level,
faking remains possible. Nonetheless, we view this as a much less
common scenario, one that requires the individual to have a basis
for perceiving that the organization views a different response to
be optimal from the one the person views to be optimal. Thus, the
first feature of interest is whether there is strong incentive for
individuals to fake when given behavioral tendency instructions
(i.e., to give knowledge should responses when given behavioral
tendency would instructions).

Factor 2 deals with whether test takers are in a familiar domain,
where there is a basis for conscious and articulated knowledge of
the domain, or in an unfamiliar domain. Familiarity with a domain
may result from specific training or experience (as in the case of an
SJT dealing with how a firefighter should deal with certain events
at a fire) or from more general life experience (as in the case of a
more general SJT dealing with effective interpersonal interaction).
We note that some SJTs are made up of a mixture of domain-
specific and domain-general items.

Combining these two features yields four scenarios. If motiva-
tion to fake is high in either a familiar or unfamiliar domain, we
would expect the same response under knowledge and behavioral
tendency instructions, because test takers give should responses
regardless of instructions. If motivation to fake is low in a familiar
domain, we would expect scores under knowledge instructions to
generally be greater than scores under behavioral tendency instruc-
tions (Nguyen et al., 2005). SJTs with knowledge instructions
would yield higher scores than SJTs with behavioral tendency
instructions unless would and should responses truly are identical
for a given candidate. However, if motivation to fake is low in an
unfamiliar domain, the situation is more complex. If the test taker
has no articulated domain knowledge and therefore does not know
the best response, then it is possible that a knowledge response
(what one cognitively assesses as the best response) is less effec-
tive than the test taker’s true (i.e., unfaked) behavioral tendency
response.

This analysis suggests that systematic differences between
knowledge and behavioral tendency instructions (i.e., higher mean
scores under knowledge instructions than under behavioral ten-
dency instructions) are likely to occur only in (a) a familiar domain
with (b) little to no motivation to fake. Of interest, we found that
the large majority of the research supporting differences between
knowledge and behavioral tendency instructions came from just
these scenarios. McDaniel et al.’s (2007) meta-analysis, which
documented a difference between knowledge and behavioral ten-
dency instructions, focused almost entirely on studies of job in-
cumbents, who are typically in a familiar domain and have little
motivation to fake.

In the present study, applicants for medical school did have a
clear motivation to obtain high scores. The domain of the SJT,
namely interpersonal skills, is one that is not explicitly taught but
rather is developed through life experiences. Thus, we expected
that applicants who were motivated to obtain high scores would be
in a position to improve their scores by giving a knowledge
response even if under behavioral tendency instructions. As evi-
dence that the domain being tested with this SJT was familiar to
applicants, we had psychology students take the SJT with no clear
incentive to fake. Scores were 0.5 SD higher under knowledge
instructions than under behavioral tendency instructions (details
are described later).

Thus, the present setting fits in the motivation to fake/familiar
domain quadrant of our $2 \times 2$ framework. As noted earlier, under
such conditions, we hypothesized no meaningful difference be-
tween knowledge and behavioral tendency instructions, keeping
SJT content constant (Hypothesis 1). We note that we define
meaningful in terms of effect size, rather than statistical signifi-
cance, because the large sample size in this study results in very
small effects reaching statistical significance (i.e., a difference
accounting for 0.2% of variance would reach statistical signifi-
cance). Using Cohen’s (1988) rule of thumb defining $d = .20$ (i.e.,
accounting for 1% of variance) as a small effect, we hypothesized
that the difference between the two types of instructions would not
reach the threshold of a small effect.

Turning to the question of the cognitive loading of the SJT, the
logic of knowledge instructions creating a maximum performance
setting suggests that knowledge instructions would maximize the
cognitive loading of a given set of SJT items. Without incentive to
fake, behavioral tendency instructions would be expected to pro-
duce a lower cognitive loading, because the responses would
reflect typical performance: what one would do rather than knowl-
dge of what one should do (McDaniel et al., 2007). We did expect
some cognitive loading under both instructional sets, because
individuals higher in cognitive ability should have been better able
to assess the effectiveness of various courses of action and should
have developed their behavioral tendencies accordingly. With in-
centive to fake in a high-stakes context, however, as in the present
applicant setting, responses under behavioral tendency instructions
should have had increasing cognitive loading as the number of
items on which the applicant gave knowledge responses increased,
and could have had a cognitive loading equal to that for knowledge
instructions if applicants gave knowledge responses to a prepon-
derance of items. The faking literature, however, shows variation
in the degree to which individuals distort responses, even when
encouraged to do so (McFarland & Ryan, 2000). Thus, consistent with past research, we hypothesized that in a high-stakes selection context, the correlation between scores on an SJT with knowledge instructions and a cognitive ability measure would be meaningfully higher than the correlation between scores on an SJT with behavioral tendency instructions and a cognitive ability measure, keeping SJT content constant (Hypothesis 2). In keeping with the reliance on practical significance necessitated by the large number of participants, as in the case of Hypothesis 1, we defined meaningfully higher as exceeding the threshold for a small effect (i.e., an effect accounting for at least an additional 1% of variance).

Finally, although we examined the criterion-related validity of an SJT under behavioral tendency and knowledge instructions, we do not offer a specific hypothesis. Prior research has produced competing findings: One can make conceptual arguments in both directions (i.e., knowledge instructions would yield higher validity, given the strong record of knowledge tests in general, vs. behavioral tendency instructions would yield higher validity because they better capture behavioral intentions).

Method

Sample

The study was situated within the context of admission to college for medical studies. Data were collected during the admission exam for medical studies administered in Belgium. The total sample consisted of 2,184 candidates (772 men and 1,412 women; 99.2% White). The average age of the candidates was 18 years and 7 months.

Students were randomly assigned to two conditions: 1,086 students (67.1% female and 32.9% male; 99.4% White; mean age = 18.7 years) completed the SJT with knowledge instructions, whereas the other 1,098 students (62.2% female and 37.8% male; 99.1% White; mean age = 18.7 years) completed exactly the same SJT but with behavioral tendency response instructions.

Procedure

The admission exam, of which the SJT was a part, lasted for a whole day and was centrally administered in a large hall. The administration of the exam was highly standardized because it was guided by a minute-by-minute script. On average, the passing rate of the admission exam was between 25% and 30%.

A week after the exam, candidates obtained feedback on their test scores. Candidates who passed received a certificate that warranted entry into any Belgian medical university. Thus, there was no further selection on the part of universities. However, not all students who passed the exam chose to enter medical studies.

Development of the SJT

The written SJT used in this study measured interpersonal/communication skills (i.e., skills other than cognitive ability) related to the interaction between a physician and a patient. We used an approach analogous to that used in other studies (Weekley, Ployhart, & Holtz, 2006) for developing the SJT. First, research assistants interviewed 10 experienced physicians and professors in general medicine (8 men, 2 women; average age = 39.5 years; average years of experience = 12.5 years) to collect critical incidents related to the domain of interest (i.e., interpersonal/communication skills related to the interaction between a physician and a patient). Research assistants familiar with SJT development then used these critical incidents to construct item stems. Next, another group of subject matter experts were asked to generate response alternatives. Accordingly, a large number of SJT items were created. Finally, for the scoring key, 10 subject matter experts (experienced physicians, professors in general medicine) independently completed all items. Agreement among the experts was generally satisfactory (Cohen’s k > .70), and discrepancies were resolved through discussion, leading to the scoring rule. In some cases, it was necessary to change or remove the items or options and insert new ones. This scoring rule indicated which response alternative was optimal for a given situational item. Endorsement of this response alternative gave the student one point. It was forbidden by law to use different scoring rules. Prior research has shown that the test has adequate criterion-related validity (Lievens & Sackett, 2007).

The final written SJT consisted of short descriptions of key interpersonal situations that physicians were likely to encounter with patients. The language of the SJT was Dutch. In total, the SJT consisted of 30 multiple-choice questions, with four response alternatives each. The testing time of the SJT was 40 min. As noted earlier, students were randomly assigned to one of two conditions. In the two conditions, the SJT had the same content (i.e., the verbal content of the situations and the response alternatives was held constant), but the 30 situations had different response instructions. In one condition, each SJT situation ended with the following response instruction: “Pick the best response” (knowledge instruction). In the other condition, each situation ended with the following question: “What would you most likely do?” (behavioral tendency instruction).

The internal consistency coefficients of the SJT across conditions were similar (α = .55 for knowledge-based instructions; α = .56 for behavioral tendency instructions). Prior research obtained with an alternate version of the SJT used in this study found a test–retest reliability of .66 (Lievens, Buyse, & Sackett, 2005b).

This specific SJT and its two conditions were tested in a low-stakes context among 83 psychology undergraduate students in a large Belgian university (68 female, 15 male; mean age = 21.4 years). These students were randomly assigned to a knowledge or behavioral tendency condition and completed the SJT for research purposes. Participants in the knowledge instruction condition (M = 15.36, SD = 3.25) scored significantly better than participants in the behavioral tendency instruction condition (M = 13.77, SD = 3.37), t(81) = 2.19, p < .05; d = .48. These results are consistent with prior research that has been conducted predominantly in low-stakes contexts (Nguyen et al., 2005). They also show that our translation of the instruction sets of prior research from English to Dutch was adequate.

Other Predictor Measures

Cognitive ability test. This test contained 50 items with five response alternatives each. The items were formulated in verbal, numeric, or figural terms. Hence, this was a broad cognitive ability test that aimed to measure general mental ability. In light of test security, we cannot mention the source of this cognitive ability test. For the same reasons, we cannot present sample items. Inter-
ested researchers may contact us to obtain more information. Prior research has attested to the good reliability and predictive validity of this test for a medical student population (Minnaert, 1996). In particular, Minnaert (1996) reported an internal consistency of .84 and a validity coefficient of .36 for predicting first-year GPA in medical studies. In this study, the internal consistency coefficient equaled .80.

Operational composite score. In addition to the SJT and the cognitive ability test, the admission exam consisted of science tests and a work sample (silent reading protocol). A weighted sum of all tests used (i.e., operational composite score) was computed to make admission decisions. Next, a minimal cutoff was determined on this composite. The weights and cutoff score were determined by law.

Criterion Measure

Criterion data were gathered from students who had passed the exam and had completed the first year of medical studies in Belgium. We retrieved archival data on students’ scores on interpersonally oriented courses at the end of the first year from all Belgian universities. These courses focused on teaching interpersonal and communication skills. Hence, they typically consisted of interactional exercises and exams. These criterion data were gathered because they were especially useful for validating an SJT that aimed to measure interpersonal and communication skills (see Lievens, Buyse, & Sackett, 2005a). A composite score (hereinafter called interpersonal GPA) was obtained by averaging scores on interpersonal courses. Interpersonal GPA correlated .38 with GPA. Prior research showed that this measure had adequate reliability (in the form of temporal stability; Lievens et al., 2005a).

Results

Table 1 presents the means and standard deviations on the cognitive ability test, the SJT, and the operational composite score in the two conditions. We investigated Hypothesis 1 at both the mean score and the item score level. In terms of mean score comparison, we found a statistically significant difference between the two conditions, with the SJT with behavioral tendency instructions (\(M = 14.92, SD = 4.00\)) receiving somewhat higher scores than the SJT with knowledge instructions (\(M = 14.52, SD = 3.95\)). A paired samples t-test revealed a statistically significant difference between the two conditions, \(t(2182) = 2.33, p < .05; d = .10\). This difference is well below the \(d = .20\) threshold for a small effect; thus, these results support Hypothesis 1. Additional analyses wherein we controlled for gender, nationality, and age confirmed these findings. At the item level, in 16 cases the behavioral tendency scores were higher than the knowledge scores, whereas the knowledge scores were higher for 14 items.

Table 2 presents the correlations between the SJT and the cognitive ability test, as the basis for testing Hypothesis 2. The correlation between the SJT with knowledge instructions and cognitive ability (\(r = .19\)) was significantly higher (\(z = 1.91; p < .05\)) than the correlation between the SJT with behavioral tendency instructions and cognitive ability (\(r = .11\)). The SJT–ability correlations accounted for an additional 2.4% of variance in the knowledge condition, thus exceeding the threshold for a small effect and supporting Hypothesis 2.

We also examined the criterion-related validity of the SJT under behavioral tendency and knowledge instructions. To deal with indirect range restriction (only students who scored higher than the cutoff determined on the operational composite passed), we applied the multivariate range restriction formulas of Reu, Carretta, Earles, and Albert (1994) to the uncorrected correlation matrix. Statistical significance was determined prior to correcting the correlations (Sackett & Yang, 2000). The correlation between the SJT with knowledge instructions and interpersonal GPA (uncorrected \(r = .15\), corrected \(r = .18\)) was not significantly different (\(z = -0.48, ns\)) from the correlation between the SJT with behavioral tendency instructions and interpersonal GPA (uncorrected \(r = .17\), corrected \(r = .19\)). A regression analysis in which data across the two response instruction conditions were pooled confirmed these results. In this regression, we first entered the cognitive ability test, a dummy representing the type of response instruction, and the SJT. Next, we entered the interaction term between the SJT and the response instruction dummy. This interaction term was not significant, and the \(R^2\) change was .00. All of this indicates that the relationship between SJT and interpersonal GPA was not statistically different from one instructional set to the other.

Discussion

This study added an important missing piece to the knowledge of response instructions and SJTs. We clarified that the effects of SJT response instructions were primarily based on studies in which there was little or no incentive to fake (incumbent settings) and in which there was a high degree of familiarity with the domain of interest. Next, we argued that SJT response instruction results might be different in high-stakes scenarios. We had the unique opportunity to conduct a field experiment wherein actual applicants were randomly assigned to two different response instruction sets. This field experiment generated the following key contributions with regard to SJT response instructions in high-stakes selection contexts.

Essentially, the answer to our initial question of whether test takers would respond differently to knowledge versus behavioral tendency instructions for SJTs in a high-stakes context seems to be no. First, there were no practically important differences between mean scores on an SJT with knowledge instructions and an SJT with behavioral tendency instructions in a high-stakes context. Therefore, the mean score differences found in prior research were not replicated in a high-stakes admission context. The high moti-
The following limitations should be noted. Our study dealt with an SJT that measured interpersonal skills. Future research should examine whether our results generalize to employment settings with job performance as the criterion.

Future research should also explore what is going on in the heads of participants who are completing SJTs with different response instructions (Ployhart, 2006; Schmitt & Chan, 2006). More generally, this study fits into a broader trend of distinguishing constructs from methods (Arthur & Villado, 2008). In this study, content was held constant, whereas the method was manipulated. Different response instruction methods were related to differential relationships with key constructs, such as cognitive ability. We need more studies that examine the isolated impact of method factors on the substantive relationships of SJT scores.

Table 2

<table>
<thead>
<tr>
<th>Predictor or criterion</th>
<th>N</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predictor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Cognitive ability</td>
<td>1,086</td>
<td>—</td>
<td>.11**</td>
<td>.64**</td>
<td>—</td>
<td>.02 (.03)</td>
</tr>
<tr>
<td>2. Situational judgment test</td>
<td>1,086</td>
<td>.19**</td>
<td>—</td>
<td>.36**</td>
<td>.17 (.19)</td>
<td>1,098</td>
</tr>
<tr>
<td>3. Operational composite</td>
<td>1,086</td>
<td>.67**</td>
<td>—</td>
<td>.40**</td>
<td>—</td>
<td>.17 (.21)</td>
</tr>
<tr>
<td>Criterion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Interpersonal GPA</td>
<td>183</td>
<td>.04 (.10)</td>
<td>.15 (.18)</td>
<td>.17 (.22)</td>
<td>—</td>
<td>168</td>
</tr>
</tbody>
</table>

Note. Results for knowledge instructions are below the diagonal; results for behavioral tendency instructions are above the diagonal. Corrected correlations are in parentheses. Correlations were corrected for multivariate range restriction. Statistical significance was determined prior to correcting the correlations. GPA = grade point average.

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

References

replication study in higher education. Unpublished doctoral dissertation, University of Louvain, Belgium.


Received June 5, 2008
Revision received October 20, 2008
Accepted October 22, 2008

---

**E-Mail Notification of Your Latest Issue Online!**

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at http://notify.apa.org/ and you will be notified by e-mail when issues of interest to you become available!