The Impact of Making Targeted Dimensions Transparent on Relations With Typical Performance Predictors

Kimberly A. Smith-Jentsch *

* University of Central Florida,
The research presented here investigated the impact of making targeted dimensions transparent to participants prior to their performance of a simulation exercise, on the level of dimension ratings and their correlations with typical performance predictors. Results from two studies, both employing between-subjects designs, showed that conceptually matched typical performance predictors were more positively associated with dimension ratings when targeted dimensions were not made transparent than when they were. In addition, only when targeted dimensions were not made transparent did conceptually matched typical performance predictors correlate more positively with dimension ratings than conceptually distinct typical performance predictors. Finally, those who were made aware of targeted dimensions received higher mean ratings in Study 1 but not in Study 2.

Assessment centers (ACs) and work simulations are often used as predictors of typical performance on the job (Thornton & Byham, 1982). In addition, they are sometimes used as typical performance criteria for the purpose of training validation (e.g., Smith-Jentsch, Salas, & Brannick, 2001). However, some would argue that performance in such exercises better represents maximum performance than typical performance. Typical performance in its purest sense reflects what a person “will do” on the job over a sustained period of time, when they are unaware that their performance is being evaluated. By contrast, maximum performance reflects what a person “can do” when they are explicitly aware that they are being evaluated, accept instructions to maximize effort, and are required to perform for a short-enough period of time that their attention remains focused on the task (Sackett, Zedeck, & Fogli, 1988). It is important to note, however, that most au-
Authors in this area agree the maximum–typical performance distinction is more appropriately viewed as a continuum than a dichotomy with situations being arrayed as to the degree to which individual differences in noncognitive factors affect performance.

Critics of the AC method argue that it produces ratings that simply represent maximum performance and does not add incremental predictive value over measures of cognitive ability (e.g., Schmidt & Hunter, 1998). Others argue that the task complexity and length of behavioral exercises can and do enable one to sample typical behavior (cf. Sackett et al., 1988). Unfortunately, little support has been shown for the notion that AC ratings reflect consistent and distinct traits (e.g., Bycio, Alveres, & Hahn, 1987; Russell, 1987; Sackett & Dreher, 1982). This lack of support is particularly problematic if such ratings are to be used to provide trait-based developmental feedback or to determine remedial training needs (Sackett & Dreher, 1984). Thus, many researchers have explored factors that may enhance the construct validity of AC ratings, noting great variability in the manner that they are implemented (cf. Lievens, 1998). One such variable is skill transparency.

SKILL TRANSPARENCY

Kleinmann (1993) defined skill transparency as the degree to which respondents correctly recognize a targeted dimension and the behaviors that compose it. Results from a survey conducted by Spychalski, Quinones, Gaugler, and Pohley (1997) suggest that 30% of AC’s provided information to participants intended to make targeted dimensions transparent. Despite this, only a small number of studies have tested the effects of making targeted dimensions transparent (Kleinmann, 1993, 1997; Kleinmann, Kuptsch, & Koller, 1996; Kolk, Born, & van der Flier, 2003; Smith-Jentsch & Marion-Landis, 1993; Smith-Jentsch et al., 2001). In the first of these studies, Kleinmann (1993) demonstrated that a substantial number of participants failed to correctly recognize targeted dimensions without being so informed. Kleinmann also found that participants tended to be rated higher on a particular dimension in exercises where they correctly recognized it than they did in exercises where they did not. These findings have since spawned a debate regarding whether targeted dimensions should be made transparent for all participants (Kleinmann, 1997; Kleinmann et al., 1996; Kolk et al., 2003; Smith-Jentsch & Marion-Landis, 1993; Smith-Jentsch et al., 2001). On one hand, if the ability to recognize targeted skills results from unequal access to assessment information, then making targeted skills transparent for all would have the effect of removing sources of contamination and “leveling the playing field.” However, if the same skills are used to determine appropriate behavior in an exercise and on the job, making targeted skills transparent should reduce criterion-related validity. At least one study demonstrated support for this notion (Smith-Jentsch et al., 2001). The
two studies reported here build on this prior research by testing the proposition that making targeted dimensions transparent elevates ratings and reduces the convergent and discriminant validity of those ratings with respect to “typical performance” predictors.

MAXIMUM AND TYPICAL PERFORMANCE PREDICTORS

The maximum and typical performance distinction has been employed with respect to predictor variables as well as performance criteria (Sackett et al., 1988). Maximum performance predictors measure the underlying ability-related constructs (e.g., cognitive ability, psychomotor skill) that enable individuals to be successful on maximum performance criteria. By contrast, typical performance predictors measure nonability constructs such as personality or goal orientation that bridge the gap between what individuals can do and what they typically will do. At least one study has shown that maximum performance predictors are more strongly related to maximum performance than to typical performance criteria (Dubois, Sackett, Zedeck, & Fogli, 1993). Conversely, the relationships between typical performance predictors and typical performance criteria have been shown to be greater for dimensions of performance with a strong nonability component (e.g., personality) than for those primarily associated with ability (Bartram, 2005).

If making targeted skills transparent removes performance variability associated with noncognitive factors, this should increase correlations between dimension ratings and maximum performance predictors. Consistent with this notion, Kolk et al. (2003) found that verbal intelligence was a better predictor of dimension ratings for a group of participants who were made aware of targeted dimensions than it was for a second group who were not. The logical converse of this would be that typical performance predictors should be more strongly associated with criterion performance ratings when targeted skills are not made transparent. However, this presumes that the variance removed by skill transparency is trait relevant. Instead, it has been argued by some that when targeted skills are not made transparent, this simply favors those with the cognitive ability to correctly recognize the behavioral requirements of the exercise. In fact, there is some evidence to suggest that individuals who score higher on general abilities such as verbal intelligence and social judgment competence are consistently better at recognizing various targeted dimensions across different types of exercises (Kleinmann, 1997; Kolk et al., 2003). If these same general abilities are used to determine what is most appropriate on the job as well, this alone could explain why making targeted skills transparent seems to reduce criterion-related validity. However, this would also mean that dimension ratings should not be treated as if they represented the specific traits they were intended to measure for the purpose of feedback or training.
needs diagnosis. Thus, it is important to understand what type of variance is removed from performance when participants no longer have to determine for themselves how best to direct their efforts. Toward this goal, the two studies reported here compared relationships between conceptually matched and conceptually distinct personality measures and criterion dimension ratings when targeted dimensions were made transparent and when they were not.

**CONSTRUCT VALIDITY**

Construct validity has been defined as the extent to which a measure reflects an underlying concept or trait. It pertains to the meaning of scores. In other words, why does one individual score higher and another lower on a particular measure? The construct validity of a measure is commonly inferred from the pattern of correlations between that measure and measures of other constructs that are theoretically related or unrelated within a nomological network. The construct validity of AC ratings has been most commonly tested by comparing correlations of same-dimension ratings and different-dimension ratings within and between multiple AC exercises. Using a similar procedure, prior research has found that correlations between same-dimension ratings across exercises (convergent validity) increase, and correlations between different-dimension ratings within an exercise decrease (discriminant validity) when targeted dimensions are made transparent in both, as compared to when they are not made transparent in either (Kleinmann, Kuptsch, and Koller, 1996, 1997; Kolk et al., 2003). On this basis, it has been argued that making targeted skills transparent increases the construct validity of dimension ratings. I argue, more specifically, that making targeted skills transparent may result in a more construct valid measure of the ability to demonstrate a particular skill but a less construct valid measure of the underlying traits that motivate the use of that skill in typical performance situations. In other words, I suggest that when individuals must determine for themselves what the most effective behavioral strategy is in a behavioral assessment exercise (i.e., one that will allow them to receive higher ratings), they rely on their own traits to guide them.

Motowidlo, Hooper, and Jackson (2006) offered the notion that personality traits have causal effects on implicit beliefs about the importance of those traits for behavioral effectiveness. These authors suggested that “when a problematic situation demands an expression of a particular trait for effective resolution, people who possess that trait are more likely to believe that behaviors expressing that trait will be effective in that situation” (p. 751). Motowidlo et al. tested this notion in a series of studies involving situational judgment tests (SJTs). Participants were asked to rate the effectiveness of multiple options for responding to a problematic work situation. The responses were designed so as to vary in the degree to which they depicted expressions of the desired traits (e.g., extraversion, agreeableness). These ratings were then used
to calculate participants’ “implicit trait policies” (IPTs). Results indicated that participants’ scores on personality measures of the desired traits were significantly correlated with their IPTs regarding behavioral expressions of those traits. Similarly, I argue that when targeted skills are not explicitly made transparent to participants before a behavioral assessment exercise, those possessing the desired traits will be more likely to accurately determine when behavioral expressions of those traits are appropriate and will direct their energy toward demonstrating those behaviors to the best of their abilities. As a result, dimension ratings will contain variability attributable to both trait and ability. By contrast, making dimension ratings transparent will lead to ratings that merely reflect ability to demonstrate the desired behaviors (i.e., one’s maximum performance).

**STUDY 1**

The trait targeted in this research was assertiveness. Assertiveness has been defined as the ability and disposition to state one’s opinions, ideas, concerns, and desires in a manner that is direct and to the point without being offensive, demeaning, or hostile. This trait has long been linked to the psychological well-being of individuals in the clinical literature (Eisler, Miller, & Hershen, 1973) and, more recently, has been recognized as important for effective team decision making (Jentsch & Smith-Jentsch, 2001). Prior research has indicated that assertiveness is multidimensional (e.g., Lorr & More, 1980). In this regard, four relatively independent dimensions of behavior have emerged from factor analyses of self-report assertiveness measures: (a) directiveness (e.g., initiative, leadership), (b) social assertiveness (e.g., initiating social relationships), (c) defense of interests (e.g., refusal of unreasonable requests), and (d) independence (e.g., maintaining one’s opinions when faced with opposition). Moreover, self-report measures of the four dimensions have shown differential predictive validities, reflecting their conceptual match to behaviors assessed in a simulation exercise (Smith-Jentsch, Salas, & Baker, 1996). Consistent with the arguments presented earlier, I expected that a conceptually matched self-report measure of the targeted assertiveness dimension would be a better predictor of criterion dimension ratings if participants were not informed of the targeted dimension than if they were. My first hypothesis stated the following:

**H1:** Ratings of directiveness in a simulation exercise will be more positively related to a self-report measure of directiveness for those in the non-transparent condition than for those in the transparent condition.

Although I expected that the self-report measure, being a typical performance predictor, would correlate less strongly with dimension ratings in the transparent
condition, I did not necessarily expect them to be uncorrelated. This is because nonability constructs can be correlated with ability-related constructs. In some cases, abilities may lead to dispositions. For instance, cognitive ability tends to correlate positively with self-efficacy (Chen, Casper, & Cortina, 2001). In other cases, personality traits appear to drive individuals to seek out certain types of experiences, and those experiences in turn lead them to develop enhanced abilities. For instance, those who score higher on openness to experience tend to seek out opportunities to learn and also tend to learn more from training (Barrick & Mount, 1991; Lievens, Harris, Van Keer & Bisqueret, 2003). Using the example of assertiveness, both those with a disposition toward defending their rights and those with a disposition toward directing others are likely to have created many opportunities to learn through trial and error how to communicate their concerns in a direct yet nonconfrontational manner. In this regard, those scoring highly on either one of these dispositions may have developed the fundamental skills required to perform equally well under the demand characteristics imposed by skill transparency although they may differ in their willingness to apply this skill in typical performance situations, depending on the purpose of communication (e.g., to defend oneself or to take control of a meeting) or the characteristics of the target (e.g., family member, stranger). As such, the differential predictive validity of conceptually matched and conceptually distinct self-report measures should be more pronounced when the targeted dimension is not made transparent. Hypotheses 2 and 3 are as follows:

H2: Self-reported directiveness will be more positively related to ratings of directiveness in the simulation than will self-reported defense of interests for those in the nontransparent condition.

H3: The relative superiority of self-reported directiveness over self-reported defense of interests as a positive predictor of directiveness behavior ratings will be greater for those in the nontransparent condition than for those in the transparent condition.

In a selection- or promotion-related assessment situation, most participants are motivated to demonstrate whatever dimensions they believe are being targeted to the best of their ability. When targeted skills are not made uniformly transparent, however, some participants who have the requisite ability to perform on those dimensions will misdirect their effort toward demonstrating performance on the wrong dimension to the detriment of the one being targeted. The resulting ratings would be an underestimate of maximum performance for those individuals. Theoretically, making targeted skills transparent should eliminate or substantially reduce such misdirection of effort and, as a result, ensure that dimension ratings better reflect “maximum performance.” However, in a study by Kleinmann (1997) and two by Kolk et al. (2003), those for whom targeted skills were made transpar-
ent did not show elevated ratings when compared to those who were not so informed. One possible reason for the lack of expected between-group differences may be that making targeted skills transparent places demand characteristics on participants which motivate them to engage in impression management. Impression management has been defined as individuals’ attempts to control the images that they project in social interactions (Schlenker, 1980). Of particular relevance to the current discussion is the aspect of impression management involving opinion conformity, or “expressions of beliefs, values, or attitudes that are known or can be reasonably assumed to be held by the target to target organization” (McFarland, Ryan, & Kriska, 2003, p. 643). Prior research has shown that the use of impression management tactics elevated ratings from a structured interview but was not related to ratings based on a role-play exercise (McFarland et al., 2003). In this regard, it was noted that structured interviews simply require candidates to describe how they would deal with a particular situation, whereas behavioral exercises (e.g., role-plays, simulations) require that candidates actually demonstrate the manner in which they would deal with a situation. Although individuals can demonstrate skills that they do not typically use, they cannot “fake” a skill that they do not have. In fact, prior research has demonstrated that for some individuals (i.e., low self-monitors), the use of impression management tactics actually had a negative impact on the manner in which they were perceived by others (Turnley & Bolino, 2001). Moreover, because impression management tactics deplete cognitive resources (Kanfer & Ackerman, 1989), their use may cause participants to underperform on a skill they actually possess. This effect may be exacerbated when multiple skills are made transparent at once, as in both Kolk et al. (2003) and Kleinmann (1997). In such cases, performance on one of the targeted skills may be hindered by the cognitive demands of impression management on a second or third skill. Thus, the expected elevation of ratings due to skill transparency is most likely to appear when a single dimension is being targeted. Therefore, my final hypothesis for Study 1 is as follows:

**H4:** Participants for whom a single targeted dimension (i.e., directiveness) is made transparent prior to performance in a behavioral simulation will be evaluated higher on that dimension than will those for whom the targeted dimension is not made transparent.

**Method**

**Participants and simulation.** Eighty-four undergraduate psychology students were recruited to participate in a 15-min simulation and received experimental credit for their cooperation. Embedded within the simulation were five trigger events scripted to occur within the context of a helicopter mission. Participants played the role of copilot, and a live confederate played the role of captain. The
confederate was tightly scripted to ensure that all participants had the same opportunity and need to demonstrate directiveness. For instance, at one point in the scenario the captain was scripted to become lost. This was made obvious to the copilot, who then had the opportunity to take the initiative to get them back on course.

Procedure. First, participants were recruited to complete self-report measures of directiveness and defense of interests in their undergraduate classes. All were specifically instructed to respond to the items as honestly as possible and were given no motivation to do otherwise. Approximately 2 weeks later, a second researcher recruited individuals from these same classes to participate in the simulation exercise. The connection between the two events was concealed so as to avoid influencing skill transparency for those in the condition where participants were not informed of the targeted dimension. Participants were randomly assigned to one of two experimental conditions. In the first condition, a researcher explicitly informed them prior to performing the simulation exercise that their assertive communication skills would be assessed. The specific behaviors that define directiveness were described, and examples of effective and ineffective responses were provided. Finally, participants were instructed to display behaviors consistent with directiveness to the best of their ability. In the second condition, participants were simply told that their teamwork skills would be evaluated and were instructed to coordinate with their teammate to the best of their ability. Participants in this condition were led to believe that their teammate was also a participant in the experiment.

Measures. Participants completed the Directiveness and Defense of Interests subscales of the assertiveness inventory by Lorr and More (1980). Participants responded to each item on a 6-point Likert scale according to how characteristic a behavioral response was of them (e.g., “I nearly always argue for my viewpoint if I think I am right”; “In an emergency, I get people organized and take charge”). Internal consistency for the eight items contained in the Directiveness subscale was estimated using coefficient alpha (α = .71), as was the internal consistency for the eight items representing Defense of Interests (α = .70).

Two condition-blind raters evaluated each participant’s use of behaviors consistent with directiveness from audiotapes made during the simulation. Five-point Likert-type scales were used. The average correlation between the two raters’ ratings across the five events ranged from .63 to .92. The two raters’ ratings for each event were then averaged to form five average event scores. The internal consistency of these five average event ratings using coefficient alpha was estimated to be .76, indicating that participants’ behavior was relatively consistent. Thus, the mean of the five average event ratings was used as an overall behavioral rating of the dimension directiveness.
Results

The data were analyzed using SPSS for Windows, Version 12.5, and the alpha level was set at .05, unless specified otherwise. As shown in Table 1, the correlation between scores of self-reported directiveness and behavioral ratings of directiveness was statistically significant for the group of participants who were not informed of the targeted skill \((r = .52, n = 40, p < .01)\) but not for the group of participants who were \((r = .10, n = 44, p > .05)\). In support of Hypothesis 1, these two coefficients were significantly different \((z = 2.10, p < .05,\) one-tailed). These results indicated that convergent validity with respect to relations among dimension ratings and self-reported personality was hindered by making targeted dimensions transparent. With respect to discriminant validity, self-reported directiveness was a significantly better predictor of directiveness ratings \((r_{\text{direct}} = .52)\) than was self-reported defense of interests \((r_{\text{doi}} = .08)\) in the nontransparent group \((\text{Steiger’s } Z = 2.58, p < .01)\). Thus, Hypothesis 2 was also supported. In the transparent condition, however, the difference between predictive validities for the two self-report measures \((r_{\text{direct}} = .10, r_{\text{doi}} = .18)\) was not significant \((\text{Steiger’s } Z = 0.44, ns)\).

To test Hypothesis 3, I compared the two absolute differences between the correlation coefficients in the transparent and nontransparent samples, using overlap between the confidence intervals around the point estimates as the statistical test. I followed the procedure suggested by Payton, Greenstone, and Schenker (2003), who argued that “using 83% or 84% size for the [confidence] intervals will give an approximate \(\alpha = 0.05\) test” (p. 38). Consequently, I compared the 83% confidence intervals around the difference in \(r\) in the two samples for overlap. In support of Hypothesis 3, the difference between the predictive validities of the Directiveness and Defense of Interests subscales was significantly larger \((d(r)1 = 0.44, CI_{83} = [0.28; 0.61])\) in the nontransparent condition than it was in the transparent condi-

### Table 1

<table>
<thead>
<tr>
<th>Targeted Dimension</th>
<th>Variable ID</th>
<th>(M)</th>
<th>(SD)</th>
<th>(V1)</th>
<th>(V2)</th>
<th>(V3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Not made transparent</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Directiveness ratings</td>
<td>V1</td>
<td>2.48</td>
<td>0.93</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Self-reported directiveness</td>
<td>V2</td>
<td>4.09</td>
<td>0.82</td>
<td>.52***</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Self-reported defense of interests</td>
<td>V3</td>
<td>4.38</td>
<td>0.69</td>
<td>.08</td>
<td>.38**</td>
<td>—</td>
</tr>
<tr>
<td><strong>Made transparent</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Directiveness ratings</td>
<td>V1</td>
<td>3.00</td>
<td>1.08</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Self-reported directiveness</td>
<td>V2</td>
<td>4.16</td>
<td>0.71</td>
<td>.10</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Self-reported defense of interests</td>
<td>V3</td>
<td>4.59</td>
<td>0.70</td>
<td>.18</td>
<td>.29**</td>
<td>—</td>
</tr>
</tbody>
</table>

\(a_n = 40, b_n = 44.\)

\(^*p < .05\) (directional test). \(^{**}p < .05\) (nondirectional). \(^{***}p < .01\) (nondirectional).
tion \( d(r)^2 = 0.08, CI_{83} = [-0.12; +0.27] \). This suggests that discriminant validity with respect to typical performance predictors was hindered by making the targeted dimension transparent as well.

Finally, in support of Hypothesis 4, a \( t \) test revealed that the group of participants for whom the targeted skill was made transparent received higher dimension ratings of directiveness in the simulation exercise \( (n = 44, M = 3.0, SD = 1.08) \) than did those for whom the targeted skill was not made transparent \( (n = 40, M = 2.48, SD = .93) \), \( t(82) = 2.37, p < .05 \).

**Discussion**

Predictive validities for conceptually matched and conceptually distinct self-report measures of the relevant trait (i.e., assertiveness) indicated both greater convergent and discriminant validity for ratings in the nontransparent sample than in the transparent sample. In addition, participants for whom the targeted dimension (i.e., directiveness) was made transparent received higher mean ratings of performance than did those for whom it was not. This finding runs counter to the assertion made by previous authors that participants with limited work experience (undergraduates) do not have the behavioral repertoires to adjust their behavior to meet the demand characteristics imposed by skill transparency (Kolk et al., 2003). Similar to Study 1, the prior research on which Kolk et al.’s assertion was made employed undergraduate students performing an individual task with a role-player that lasted, on average, 15 min. However, participants in the prior study were made aware of multiple (three or four) targeted skills, whereas in Study 1 participants were only informed of one dimension. To reconcile this apparent inconsistency in results, a second study was conducted to explore the possibility that the inconsistent effects of skill transparency on mean levels of performance were due to differences in the number of skills made transparent to participants in different studies.

**STUDY 2**

In Study 2, like in Kolk et al.’s (2003) prior research, three targeted dimensions were made transparent to participants. This time we did not expect to see a significant elevation of ratings for those in the transparent condition. However, we did expect once again that making targeted skills transparent would lower convergent and discriminant validity, as evidenced by the pattern of correlations with conceptually matched and conceptually distinct predictors. To investigate the robustness of this effect, Study 2 hypotheses were tested using dimension ratings for defense of interests rather than directiveness. Three hypotheses were offered:
H5: Self-reported defense of interests will be more positively related to dimension ratings of defense of interests in the nontransparent condition than in the transparent condition.

H6: Self-reported defense of interests will be more positively related to dimension ratings of defense of interests in the simulation than will self-reported directiveness for those in the nontransparent condition.

H7: The relative superiority of self-reported defense of interests over self-reported directiveness as a positive predictor of defense of interests behavior ratings will be greater for those in the nontransparent condition than for those in the transparent condition.

Method

Participants and simulation. Participants were 60 (36 male, 24 female) undergraduate college students (average age = 21) who performed a 40-min computer-based simulation for experimental credit. The simulation exercise was administered via a personal computer system. During the simulation, participants were presented with a video-based scenario in which characters appeared to speak directly to them. The scenario ran continuously (without pauses) for approximately 40 min. Responses were made orally into a microphone that was attached to participants’ headsets. When the computer detected a participant’s speech, it began recording his or her response; video of the character to whom the participant spoke was looped (i.e., a segment of the video repeated, creating the illusion that the character was intently waiting on a response). When the simulation no longer detected the participant’s speech for a period of 3 sec, the video ceased looping, and the simulation seamlessly transitioned to the next scene. In the scenario, participants played the role of a customer service representative in a hospital emergency room waiting area. As in Study 1, five trigger events were embedded in the simulation. This time, these events provided an opportunity for participants to refuse unreasonable requests in a manner that was direct and firm without being rude or hostile (i.e., demonstrate defense of interests). For instance, in one event, a patient asked the participant to be moved ahead of other patients on the list for the triage nurse.

Measures. Participants completed the identical self-report measures as in Study 1 for defense of interests and directiveness (Lorr & More, 1980). Behavioral ratings were based on participants’ verbal responses to each simulated event and were recorded and evaluated by a single rater using a 5-point Likert-type scale. The reliability of ratings made by the Study 2 rater had been established using a set of recordings from a previous study that a second rater had already rated. Using this sample of ratings from 99 prior participants, interrater reliability across the five trigger events ranged from .78 to .90.
Procedure. As in Study 1, participants in Study 2 completed the two self-report measures as part of a larger battery of personality measures in what appeared to them as another study, and they were not aware of the connection between the two. Again, participants were instructed to respond to items as honestly as possible and were given no incentive to do otherwise. All respondents volunteered to participate in the “second” study in which they were randomly assigned to either the transparent or nontransparent condition. The manipulation of skill transparency was identical to Study 1, with one exception. In Study 2, the ineffective examples presented included both passive and aggressive responses, whereas in Study 1, the focus was simply on avoiding passive responses. Participants in the nontransparent condition were simply told that the exercise was designed to assess interpersonal and customer service skills and were instructed to perform to the best of their ability.

Results

Descriptive statistics and intercorrelations among the Study 2 variables are listed in Table 2. As shown in this table, no significant difference existed between the mean level of dimension ratings received in the transparent ($n = 32, M = 3.26, SD = .32$) and nontransparent ($n = 28, M = 3.19, SD = .42$) conditions, $t(58) = -1.86, ns$. However, Levene’s test for equality of variances indicated that significantly greater variability existed in the ratings received by those in the nontransparent group than by those in the transparent group, $t(58) = 4.87, p < .05$. The self-report measure of defense of interests was positively and significantly correlated with ratings of defense of interests for those in the nontransparent condition ($r = .46, p < .05$). In contrast, self-reported defense of interests was negatively correlated with ratings of defense of interests for those in the transparent condition ($r = -.33, ns$).

<table>
<thead>
<tr>
<th>Targeted Dimension</th>
<th>Variable ID</th>
<th>M</th>
<th>SD</th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not made transparent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defense of interest ratings</td>
<td>V1</td>
<td>4.50</td>
<td>.64</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-reported defense of interests</td>
<td>V2</td>
<td>4.04</td>
<td>.79</td>
<td>.04</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Self-reported directiveness</td>
<td>V3</td>
<td>3.20</td>
<td>.42</td>
<td>.46**</td>
<td>.26</td>
<td></td>
</tr>
<tr>
<td>Made transparent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defense of interest ratings</td>
<td>V1</td>
<td>3.26</td>
<td>.31</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-reported directiveness</td>
<td>V2</td>
<td>4.16</td>
<td>.77</td>
<td>.26</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Self-reported defense of interests</td>
<td>V3</td>
<td>4.70</td>
<td>.83</td>
<td>-.33</td>
<td>.26</td>
<td></td>
</tr>
</tbody>
</table>

$a_n = 28, b_n = 32.
*p < .05 (directional test). **p < .05 (nondirectional). ***p < .01 (nondirectional).
The difference between these correlations was statistically significant ($z = 3.08, p < .01$), in support of Hypothesis 5.

Hypothesis 6 stated that self-reported defense of interests would be more positively related to defense of interest dimension ratings than self-reported directiveness for those in the nontransparent condition. Results were supportive of this hypothesis. Specifically, for the nontransparent group, self-reported defense of interests scores ($r_{doi} = .46$) were positively correlated with defense of interests ratings, and this correlation was significantly greater than the correlation between defense of interests ratings and self-reported directiveness ($r_{direct} = .04, ns$) (Steiger’s $Z = 1.95, p < .05$, one-tailed). In contrast, self-reported defense of interests was negatively related to dimension ratings for those in the transparent condition, whereas self-reported directiveness was positively related to defense of interests ratings. Although the difference between these predictive validities ($r_{direct} = .26$, $r_{doi} = -.33$) was also significant (Steiger’s $Z = 2.60, p < .01$), it was in the opposite direction. The confidence intervals around the absolute differences in the predictive validities (for the two self-report measures) did, albeit not by large amounts, overlap for the transparent ($d(r)_1 = 0.42, CI_{83} = [0.20; 0.61]$) and nontransparent conditions ($d(r)_2 = 0.07, CI_{83} = [-0.16; 0.30]$). However, Hypothesis 7 had stated that the relative superiority of self-reported defense of interests over self-reported directiveness as a positive predictor of defense of interests behavior ratings would be greater for those in the nontransparent condition than for those in the transparent condition. Given that self-reported defense of interests was negatively related to dimension ratings for those in the transparent condition, Hypothesis 7 was nonetheless supported.

**Discussion**

Consistent with the results from Study 1, a conceptually matched typical performance predictor (i.e., self-reported defense of interests) was more positively related to dimension ratings (of defense of interests) for the sample of participants in the nontransparent condition than for those in the transparent condition. In addition, once again, only for the nontransparent group was the conceptually matched predictor more positively associated with dimension ratings than a conceptually distinct predictor. In fact, among Study 2 participants in the transparent condition, the correlation between ratings (of defense of interests) and the conceptually matched predictor (i.e., self-reported defense of interests) actually changed sign and became negative. This may have been due to the added instructions regarding the difference between assertiveness and “aggressiveness” in Study 2. Specifically, the manipulation of skill transparency may have motivated those less inclined to defend their interests to behave more assertively than usual and at the same time motivated those strongly inclined to defend their interests to behave less assertively than usual—so as to avoid being aggressive. This would have contributed to
the fact that mean dimension ratings in Study 2 were not elevated by skill transparency as they were in Study 1. As argued earlier, the cognitive demands of self-monitoring one’s performance on multiple dimensions and engaging in impression management tactics to attempt to optimize performance ratings on all of them simultaneously may have had a negative effect on ratings for some. In addition, others may have simply misapplied the information about behavioral requirements provided to them.

GENERAL DISCUSSION

The combined results from the present and prior research (Kleinmann et al., 1993, 1996, 1997; Kolk et al., 2003) support the notion that making targeted dimensions transparent may allow for a more construct valid assessment of the abilities necessary for “maximum performance”; however, allowing skill transparency to vary naturally among participants appears to result in a more construct valid measure of traits that motivate “typical performance” of a behavioral dimension. Thus, to answer the question “Should targeted skills be made transparent?” one must carefully consider the purpose for collecting exercise ratings. If, on one hand, one desires to predict participants’ ability to perform a behavioral dimension in maximum performance situations, to determine the impact of some intervention (e.g., training) on maximum performance, or to provide feedback to trainees regarding their ability to perform at maximum capacity, making a targeted skill transparent is recommended. With respect to the concern in this case that making targeted dimensions transparent may allow participants to “fake” a characteristic, one should be concerned only if attempting to make trait-related inferences. In other words, skill transparency does not enable participants to “fake” a skill that they do not possess, although it may enable them to demonstrate a skill that they do not typically use.

On the other hand, if one desires to predict participants’ tendencies to utilize a behavioral dimension in typical performance situations or wishes to provide feedback regarding trait-based tendencies, I argue that it is best not to make targeted skills transparent. One might ask if it would not be more effective and efficient to simply use self-reported personality to account for the dispositional component of typical performance rather than a behavioral exercise. In this regard, it is important to note that participants are able to artificially elevate scores on self-report measures of personality and that this has been shown to have a detrimental effect on the factor structure of these measures (Van Iddekinge, Raymark, & Roth, 2005). In the studies presented here, participants were instructed to complete the self-report measures honestly and, unlike candidates for selection or promotion, were given no motivation to do otherwise. Thus, in an operational selection context, behavioral exercises in which targeted dimensions are not made transparent may provide
a more valid assessment of personality dispositions than more obtrusive and transparent self-report measures.

Finally, it is encouraging to note that ratings for the nontransparent group in Study 2 correlated with the relevant typical performance predictor, even though the simulation required participants to respond entirely to “canned” video-based characters and not live role-players. These types of multimedia systems are becoming increasingly popular as they reduce instructor workload and are able to depict behavioral incidents that are more realistic than do classroom role-plays (Weekley & Jones, 1997).

**Limitations, Conclusions, and Directions for Future Research**

When generalizing the results reported here, it must be noted that both studies investigated dimensions of the same overall trait (i.e., assertiveness) and utilized undergraduate college students as participants. Additional research is needed to cross-validate the findings from the research using different assessment dimensions and different organizational participants. However, the evidence to date clearly suggests that practitioners can expect the correlates of dimension ratings to vary systematically as a function of their decision to make targeted dimensions transparent or not.

**ACKNOWLEDGMENTS**

Portions of the research described in this article were presented at the 1993 meeting of the Academy of Management and the 2006 meeting of the Society for Industrial and Organizational Psychology.

The author would like to acknowledge several individuals who assisted in the collection and coding of data reported: Amy Griffin, Dana Kendall, Helen Boudreaux, Mike Flood, Moshe Feldman, Nic Bencaz, Patrick Rosopa, Shannon Scielzo, Staci Singleton, and Ugochi Onyejiaka.

**REFERENCES**


