

**Transparency of Assessment Centers:
Lower Criterion-Related Validity but Greater Opportunity
to Perform?**

Pia V. Ingold and Martin Kleinmann

Universität Zürich

Cornelius J. König

Universität des Saarlandes

Klaus G. Melchers

Universität Ulm

Currently in press in *Personnel Psychology*

Author Note

Pia Ingold and Martin Kleinmann, Psychologisches Institut, Universität Zürich, Switzerland, Cornelius König, Fachrichtung Psychologie, Universität des Saarlandes, Germany, Klaus Melchers, Institut für Psychologie und Pädagogik, Universität Ulm, Germany

The study reported in this paper was supported by a grant from the Swiss National Science Foundation (Grant 100014-124449). We thank Morena Indelicato, Tom Siebenaler, and Aline Winzeler who helped collecting the data of this study and Anne Jansen for her advice on the data collection.

Correspondence concerning this article should be addressed to Pia Ingold, Arbeits-und Organisationspsychologie, Universität Zürich, Binzmühlestrasse 14/12, CH-8050 Zürich, Switzerland. E-Mail: p.ingold@psychologie.uzh.ch

Abstract

Assessment centers (ACs) are popular selection devices in which assesseees are assessed on several dimensions during different exercises. Surveys indicate that ACs vary with regard to the transparency of the targeted dimensions and that the number of transparent ACs has increased during recent years. Furthermore, research on this design feature has put conceptual arguments forward regarding the effects of transparency on criterion-related validity, impression management, and fairness perceptions. The current study is the first to examine these effects using supervisor-rated job performance data as the criterion. We conducted simulated ACs with transparency as a between-subjects factor. The sample consisted of part-time employed participants who would soon be applying for a new job. In line with our hypothesis, results showed that ratings from an AC with nontransparent dimensions were more criterion valid than ratings from an AC with transparent dimensions. Concerning impression management, our results supported the hypothesis that transparency moderates the relationship between self-promotion and job performance, such that self-promotion in the nontransparent AC was more positively related to job performance than self-promotion in the transparent AC. The data lent no support for the hypothesis that participants' perceptions of their opportunity to perform are higher in the transparent AC.

Transparency of Assessment Centers: Greater Opportunity to
Perform but Lower Criterion-Related Validity?

Assessment centers (ACs) are popular personnel selection procedures all over the world (Krause & Gebert, 2003; Krause, Rossberger, Dowdeswell, Venter, & Joubert, 2011; Krause & Thornton, 2009; Lievens, Harris, Van Keer, & Bisqueret, 2003). In ACs, assessee's skills and abilities are assessed on predefined dimensions during several exercises (e.g., International Task Force on Assessment Center Guidelines, 2009). Surveys show that ACs differ with regard to the transparency of targeted dimensions and that the percentage of transparent ACs has increased over time (Thornton & Krause, 2009). Specifically, in some ACs, assessee's receive detailed information about the AC dimensions with behavioral examples (i.e., dimensions are transparent), whereas in other ACs, assessee's are not provided with such information (i.e., dimensions are nontransparent).

So far, it is unclear what effects dimension transparency has on key outcomes of the selection process such as the criterion-related validity of the AC, assessee's impression management (IM) behavior during the AC, and their perceptions concerning the fairness of the AC. With regard to the effects of dimension transparency on criterion-related validity, some researchers have suggested a positive effect of transparency on criterion-related validity (e.g., Dodd, 1977; Maurer, Solamon, & Lippstreu, 2008), whereas others have warned of a negative effect on criterion-related validity (e.g., Kleinmann et al., 2011; Smith-Jentsch, 2007). Furthermore, it has been argued that transparency might make it easier for, or even encourage, assessee's to show behaviors that are different from how they would typically act (Smith-Jentsch, 2007; Thornton & Gibbons, 2009). Transparency has also been argued to influence fairness perceptions (Thornton & Gibbons, 2009). In particular, transparency may affect the opportunity to perform (Gilliland, 1993), which refers to the assessee's perception that they have an opportunity to demonstrate their knowledge, skills and abilities (Schleicher,

Venkataramani, Morgeson, & Campion, 2006). When assesseees are informed about dimensions, they could perceive an increased opportunity to show the requested behavior.

In summary, despite the increasing percentage of transparent ACs and a number of arguments about effects of transparency on important selection outcomes, there is no empirical evidence on whether making dimensions transparent to assesseees affects outcomes of interest. As such, the goal of our study was to understand (a) whether transparency affects AC criterion-related validity, (b) whether transparency affects the relation between assesseees' IM behavior and job performance, and (c) whether transparency affects fairness perceptions of the AC. By exploring these questions, the present study should provide evidence-based insight on whether dimensions should be made transparent to participants in operational ACs.

Transparency of Dimensions in ACs

Transparency has been referred to as the degree to which applicants are informed about the behavioral dimensions that are being assessed in a selection procedure (e.g., Klehe, König, Richter, Kleinmann, & Melchers, 2008; Kleinmann, Kuptsch, & Köller, 1996). Accordingly, we view transparency of ACs on a continuum from very nontransparent to very transparent. A very nontransparent AC would provide no information at all about the dimensions that are being measured, whereas a very transparent AC would provide detailed information about the dimensions and behavioral examples for these dimensions before each AC exercise. According to an international survey (Thornton & Krause, 2009), 43% of ACs disclose information about the targeted dimensions, which is a higher percentage than about one decade earlier when dimensions were only made transparent in 29% of ACs (Spsychalski, Quinones, Gaugler, & Pohley, 1997). This development requires further research given that that transparency may influence validity and fairness perceptions of assesseees, which are key criteria when implementing selection devices.

Transparency and Criterion-related Validity

It has been argued that disclosing information about targeted dimensions influences AC criterion-related validity (Dodd, 1977; Klehe et al., 2008; Kleinmann, 1997; Maurer et al., 2008; Smith-Jentsch, 2007). Some researchers assume that transparency helps provide an equal chance for all assessees and reduces uncertainty across assessees (Dodd, 1977; Kurecka, Austin, Johnson, & Mendoza, 1982; Maurer et al., 2008). Accordingly, transparent selection procedures should be more criterion valid. Specifically, if transparency reduces error variance (i.e., noise) by providing equal opportunities for the performance of all assessees, the prediction of job performance should improve.

In contrast, others have suggested that transparency diminishes the criterion-related validity of ACs as it reduces variance that is shared by the predictor and the criterion (Klehe et al., 2008; Kleinmann, 1997; Lievens, Tett, & Schleicher, 2009; Smith-Jentsch, 2007). Conceptually, the argument that dimension transparency is detrimental for criterion-related validity can be grounded in the cognitive affective personality system theory (CAPS, Mischel & Shoda, 1995). CAPS theory argues that an individual's construal of situational characteristics affects the behavior shown in different situations. The underlying mechanism is that the features of a situation interact with individual difference variables and trigger affective and mental cognitive representations. Behavioral scripts are then prompted based upon the activated representation. This theory highlights two points that are relevant for understanding assessees' behavior in ACs. First, assessees can perceive characteristics of the AC situation differently, so that affective and cognitive mental representations may differ between them (Jansen, Lievens, & Kleinmann, 2011; Kleinmann et al., 2011; Melchers, Wirz, & Kleinmann, 2012). Second, situational characteristics of the AC, such as dimension transparency, affect the mental representations and thereby influence the behavioral scripts that guide assessees' behavior.

In the vast majority of situations on the job, however, employees are not informed in advance about specific relevant performance dimensions. Hence, the degree of transparency in the selection setting affects the similarity of the selection situation (i.e., the AC) and the job situation (on which evaluations of job performance are based), with lower similarity when transparency of dimensions is higher. The resulting dissimilarity in situational cues between a transparent selection context (i.e., AC) and job situations may evoke different performance behaviors. Therefore, making dimensions transparent may diminish the situational similarity of the predictor and the criterion, thereby reducing criterion-related validity (Wernimont & Campbell, 1968).

According to interactionist approaches, situational characteristics do not fully determine behavior and behavior outcomes (i.e., performance ratings), as the situational characteristics interact with the person (i.e. individual difference characteristics). Specifically, based on CAPS theory, assessees' affective and mental cognitive representations of the situation vary between individuals and accordingly, the activated behavioral scripts may also vary between persons. For instance, some assessees in transparent ACs (i.e. in assessments in which the dimensions are made transparent) might experience the assessment situation as more controllable and are consciously aware of performing well on these dimensions and thus, manage to perform especially well. In contrast, other assessees might perform worse in the same transparent AC because knowing about the dimensions may increase their test anxiety (e.g., due to lower self-efficacy or lack of experience or skills relevant to these dimensions) which impedes their performance. Hence, one cannot expect a general increase in assessees' performance given that the effects on individual performance may differ depending on the cognitive and affective representations that are influenced by the assessees' traits, skills, abilities, knowledge and experience (Mischel & Shoda, 1995). In line with this, the majority of the few available studies that investigated the effects of transparency on AC performance found no differences (Kolk, Born, & van der Flier, 2003; Schulze Versmar,

Thomas, & Kersting, 2007; Smith-Jentsch, 2007, Study 2; Strobel, Lerche, Hoppe, & Bolte, 2006). To sum up, the relationship between dimension transparency and assessee's performance in an AC is complex.

Empirical underpinnings for CAPS and the suggestion that evaluations from a transparent AC might be less valid stem from research on individuals' ability to identify evaluation criteria in nontransparent selection procedures (Kleinmann et al., 2011). Conceptually, this ability can be linked to CAPS theory as it refers to the mental cognitive representations of the selection situation that should (according to CAPS) influence the assessee's behavioral script and hence their selection performance. A consistent body of research on this ability shows that individuals who are better at discerning the evaluation criteria (i.e., the targeted dimensions) in a nontransparent selection procedure show better performance in the procedure (Griffin, 2014; Ingold, Kleinmann, König, Melchers, & Van Iddekinge, 2014; Jansen et al., 2013; Kleinmann, 1993; König, Melchers, Kleinmann, Richter, & Klehe, 2007; Melchers et al., 2009; Preckel & Schüpbach, 2005; Speer, Christiansen, König, Melchers, & Kleinmann, 2014). In addition, recent research found a) a significant correlation between the degree to which individuals were able to discern the targeted dimensions in a nontransparent AC and supervisor ratings of their job performance, and b) that the AC was no longer criterion valid after partialling out the degree to which participants were able to discern the targeted dimensions (Jansen et al., 2013). Consequently, when dimensions are made transparent, individual differences in the ability to discern nontransparent dimensions are reduced, as is the contribution of these differences to criterion-related validity.

Previous studies on transparency also suggest that transparency might have a negative impact on the validity of ACs for predicting criteria that are proxies for job performance. In one study (Kleinmann, 1997), assessee's performance in a nontransparent AC predicted their performance in a second nontransparent AC, which served as the proxy criterion, to a

significantly larger extent than did participants' performance in a transparent AC. In a second study (Smith-Jentsch, Salas, & Brannick, 2001), the correlation between ratings of only one dimension (directiveness) in one exercise (a flight simulation) and self-reported directiveness in the cockpit one year later as the proxy criterion were significantly stronger for those in the nontransparent AC than for those in the transparent AC. In contrast, related experimental research on making dimensions transparent in interviews (Klehe et al., 2008, Study 2) found no effect of transparency on predicting AC performance as a proxy criterion of job performance.

In summary, based upon CAPS theory, related research on the ability to identify criteria, and previous evidence on transparency, we make the following hypothesis:

Hypothesis 1: Ratings from a nontransparent AC (i.e., without detailed information for assesseees about the targeted dimensions) are more criterion valid with regard to predicting task performance on the job than ratings from a transparent AC (i.e., with detailed information for assesseees about targeted dimensions).

AC performance ratings capture task-related behaviors in the AC, hence assessors often rate assesseees' behavior as it relates to the tasks required in each exercise. Similarly, the criterion task performance captures task-related behaviors on the job, hence supervisors often rate subordinates' behavior as it relates to the tasks required by the job. As such, there is a conceptual overlap between predictor and criterion. OCB ratings, in contrast, capture behavior that relates to behavior outside the employees' job demands (Organ, 1988). In other words, OCB is less similar to the predictor than task performance and it is unclear to what extent AC performance and OCB share variance. Recent research seems to support the position that AC performance ratings may not predict OCB (Hoffman, 2014). Moreover, informing assesseees about the demands of the AC tasks they are facing should not alter the relation to a criterion that captures behavior outside the task demands of the job. Accordingly, we do not suppose that making dimensions transparent will alter the relation between AC performance and OCB.

However, to further understand the effects of transparency on job performance, we will conduct the same moderation analysis for OCB to contrast the results with those obtained for task performance.

Transparency and Impression Management

Considering the broader context, selection settings provide a high-stake situation that encourages assessees to make a good impression. In line with this, research has shown that IM, which is defined as “conscious or unconscious attempts to control the images that are projected in ... social interactions” (Schlenker, 1980, p. 6), is a common behavioral response in these settings (e.g., Barrick, Shaffer, & DeGrassi, 2009; Klehe, Kleinmann, Niess, & Grazi, 2014; McFarland, Ryan, & Kriska, 2003; McFarland, Yun, Harold, Viera, & Moore, 2005). There is an intensive debate on whether assessees’ IM biases selection results, and is thus a threat to criterion-related validity (e.g., Anderson, 1991; McFarland, Ryan, & Kriska, 2002), or whether IM might be beneficial on the job (e.g., Ellis, West, Ryan, & DeShon, 2002; Ingold, Kleinmann, König, & Melchers, 2014; Klehe et al., 2014; Kleinmann & Klehe, 2011).

In this regard, it is important to note that IM does not refer to faking. Rather, recent research has highlighted that IM may include truthful and less truthful statements (see Levashina, Hartwell, Morgeson, & Campion, 2014; Melchers, Ingold, Wilhelmy, & Kleinmann, in press). Faking, however, refers to deceptive distortion (Levashina & Campion, 2007). If assessees in ACs are able to show certain behaviors, even though these behaviors are less typical for the assessee, we assume that this does not necessarily have to represent faking as this may be a component of their actual behavioral repertoire. For this reason, we focused on IM in the present study. It is also worth noting here that being in a selection setting triggers a mindset in which assessees strive to make a good impression, whether they are informed about the dimensions or not.

Two broader categories of IM can be distinguished: self-promotion and ingratiation (Stevens & Kristof, 1995). Self-promotion refers to self-focused IM behavior whereby

assessees try to appear competent by promoting their abilities and accomplishments. In contrast, ingratiation refers to other-focused IM behavior, whereby individuals try to be seen as likable by flattering others or doing favors for them. The mechanism underlying the effects of self-promotion (i.e., self-focused tactics) on ratings is supposed to be different from the mechanism underlying effects of ingratiation (i.e., other-focused tactics, Kristof-Brown, Barrick, & Franke, 2002; Wayne & Kacmar, 1991; Wayne & Liden, 1995). The reasoning is that self-promotion influences performance ratings because it enables assesseees to emphasize their qualities which may lead to a raised awareness of these qualities. Other-focused tactics, however, are supposed to influence performance ratings by increasing perceptions of liking, similarity, and interpersonal attraction in others (Ellis et al., 2002; Stevens & Kristof, 1995). Accordingly, dimension information may alter the nature of self-promotion, but it should not increase or decrease how much assessors like assesseees.

According to CAPS theory, one can assume that making dimensions transparent in a selection situation represents a change of a situational feature and may thus alter self-promotion in such a way that it deviates from an individual's self-promotion in a nontransparent situation. For instance, assesseees who know about targeted dimensions can adapt the content of their self-promotion to fit the dimensions (e.g., telling about their analytical, organizational, and presentation skills when these skills are assessed). Hence, transparency may potentially provide some guidance for assesseees' self-promotion as the dimension information enables them to self-promote with regard to the targeted dimensions.

In particular, the major issue is how typical or how biased assesseees' self-promotion is in a transparent (vs. nontransparent) selection procedure. From the point of view of CAPS theory, lowering situational similarities between the selection setting and the job setting by making dimensions transparent in the AC should lead to less similar behaviors in both settings and hence, might impair criterion-related validity. Specifically, in a nontransparent AC, assesseees have no explicit information about the dimensions that are evaluated, and hence,

their self-promotion behavior should be more typical for them. Thus, assesseees should show a similar kind of self-promotion behavior on the job, and hence, their specific self-promotion behavior should represent criterion-relevant variance of the predictor that does not compromise the prediction of job performance. In a transparent AC, assesseees may use information about the dimensions to tailor their self-promotion behaviors, (e.g., by self-promoting on targeted dimensions), thereby showing self-promotion that deviates content-wise from self-promotion in a nontransparent AC. As assesseees usually do not have comparable information about specific evaluation criteria on the job, this self-promotion would not typically be shown on the job, and thus, would introduce criterion-irrelevant variance in the predictor instead of shared variance. Accordingly, we make the following prediction:

Hypothesis 2: Transparency moderates the relation between assesseees' self-promotion and job performance, such that assesseees' self-promotion in the nontransparent AC will be more positively related to supervisor-rated job performance than assesseees' self-promotion in the transparent AC.

Conversely, we do not hypothesize a moderation effect of transparency on the relation between ingratiation and job performance. The reasons for this is that knowledge about the dimensions is of limited use for assesseees' ingratiation because it does not include information that can be used to tailor ingratiation to the target (e.g., how to flatter the target). As such, no difference is expected between ingratiation in the selection situation and ingratiation on the job. We nevertheless test transparency as a moderator of the relation between ingratiation and job performance, because this allows contrasting the results with those on self-promotion, thereby adding to a conceptual understanding of the effects of transparency with regard to these two IM tactics.

Transparency and the Opportunity to Perform

According to the International Task Force on Assessment Center Guidelines (2009), AC users and designers are supposed to consider assessee's fairness perceptions. Indeed, in a recent survey (Eurich, Krause, Cigularov, & Thornton, 2009), about 32% of US companies collect information on assessee's fairness perception in ACs, and practitioners consider fairness perception when deciding about the implementation of selection procedures (König, Klehe, Berchtold, & Kleinmann, 2010).

We consider the opportunity to perform as a crucial fairness component for transparency and therefore focus on it as a first step in exploring the relation between transparency and fairness perceptions (Thornton & Gibbons, 2009). In Gilliland's (1993) fairness model, opportunity to perform is one of several formal characteristics of a selection device that should all be related to perceptions of procedural fairness. In line with this model, previous research found that the opportunity to perform is a consistent predictor of overall fairness perceptions (Hausknecht, Day, & Thomas, 2004; Schleicher et al., 2006; Truxillo & Bauer, 2011), and is as important as job-relatedness for fairness perceptions (Schleicher et al., 2006).

From assessee's perspective, transparent ACs may have potential benefits and thus, this factor should be considered in addition to concerns about criterion-related validity. Despite being in an unfamiliar testing situation, assessee in a transparent AC may feel more in control of the situation because they are given information on the targeted dimensions (see Klehe et al., 2008). This may increase perceived procedural fairness (Gilliland, 1993), or prevent negative psychological effects associated with being tested (Anderson & Goltsi, 2006). In particular, when assessee are informed about the AC dimensions, transparency may enable them to focus on performing (as argued by Kleinmann et al., 2011; Kleinmann et al., 1996) and thus increase their perceived opportunity to perform. Hence, we examine the following hypothesis:

Hypothesis 3: Assessee's perceived opportunity to perform is higher in the transparent AC than in the nontransparent AC.

Method

Participants

Assessee's. Participants were informed about the opportunity to participate as assessee's in a simulated AC. E-mails and advertisements were sent via the career services of several universities to individuals who were currently or would soon be applying for a new job. In addition, we promoted the opportunity to participate in a simulated AC at a job fair for graduates and in the two leading national newspapers. Participating in this simulated AC allowed participants to prepare for future applications by gaining first-hand experience with ACs and receiving extensive feedback on their performance at the end of the AC. A survey on the selection practices in 125 companies across different sectors showed that 50% of ACs are used to assess trainees (Schuler, Hell, Trapmann, Schaar, & Boramir, 2007). Thus, our sample represents a relevant group of AC participants in the field. Further, we only allowed for participation in the simulated selection process when participants permitted us to contact their supervisors via e-mail to collect job performance data. To ensure that the online-survey on job performance was filled in by supervisors and not by a third party or the participant herself or himself, we asked assessee's to provide the name of the supervisor and their supervisor's e-mail address so that we could screen the match. As it is common knowledge what the e-mail addresses of companies and employers normally look like (e.g., surname@company.com) and where to find the homepages, we contacted the employees if the email format deviated from this.

We excluded three persons because their supervisors reported that they could not evaluate their employees' performance well. Afterwards, the sample contained 194 assessee's (104 males and 90 females; mean age = 28.54 years, $SD = 5.42$), had an average work experience of 29.63 months ($SD = 25.55$), and worked on average 28.5 hours per week. About

38% of the participants worked in the research and education sector, 9% in the banking and insurance sector, 7% in the industrial sector, 6% in the service sector, 6% in the media and communication sector, 5% in the public sector, 5% in health services, 4% in the traffic and transportation sector, 2% in sales and distribution, and 12% in other (non-specified) sectors.

Assessors. Assessors were 73 psychology master's students (mean age = 25.48 years, $SD = 7.98$) who had participated in one day of frame-of-reference training (Roch, Woehr, Mishra, & Kieszczyńska, 2012; Woehr & Huffcutt, 1994). In this training, assessors were informed in detail about the targeted dimensions and the respective behavioral examples for good, moderate, and poor performance. Assessors also were given practice using the AC rating scales. After having been extensively informed about the exercises, the dimensions and the rating scales, all assessors practiced assigning behaviors to the respective dimensions with video-tapes from a comparable study and received extensive feedback on the correctness of their ratings.

Facilitators. Facilitators of the AC were three Industrial and Organizational Psychology master's students who were constantly supervised by the authors of this study. They were responsible for the organization of each AC day and coordinated both participants and assessors. Two of them had started working in the HR sector.

Supervisors. The supervisors were on average 45.79 years old ($SD = 10.08$), had spent 4.63 years ($SD = 1.53$) in a supervisory function, and worked in various sectors (see above). 73.4% of the supervisors were male, 26.6% were female. In both conditions, supervisors reported that they could evaluate the performance of their employees well on a scale from 1 = *badly* to 5 = *well*, with an average score of 4.49 ($SD = 0.61$) in the nontransparent condition and 4.48 ($SD = 0.67$) in the transparent condition. Assessors had no access to participants' job performance ratings and supervisors were not informed about participants' performance in the AC.

Study Design

When deciding how to collect data, our main consideration was to allow for conclusions on the effects of transparency and to minimize internal validity threats. Accordingly, we decided to manipulate transparency in a between-subjects design with a nontransparent and transparent condition that allowed for full control of the AC and the manipulation such that both conditions could be conducted identically except for the manipulation. A further consideration was to provide an ecological valid setting. Former studies with simulated ACs (e.g., Jansen et al., 2013; Kolk et al., 2003; König et al., 2007) have illustrated that participants who are about to apply for a new job typically perceive the setting as realistic, report nervousness and try to perform at their best. Based upon these considerations, we chose to collect data in a simulated AC. A further advantage of this approach was that assesseees were not motivated to distort answers on the self-report measures of self-promotion and the opportunity to perform because their answers to these measures were not relevant for an actual job offer.

The simulated selection process included four AC exercises: two group discussions and two presentation exercises. We chose these types of AC exercises as they are the two most commonly used in Western Europe (Krause & Thornton, 2009). The four exercises we used had been successfully developed and employed in other studies (Jansen et al., 2013; Wirz, Melchers, Schultheiss, & Kleinmann, 2014) and were designed to simulate a selection process for a graduate trainee position. This position was selected because it represents a realistic and attractive job for assesseees with various backgrounds and is a common position for which ACs are used (Schuler et al., 2007). In the first presentation, assesseees were instructed to present one of their leisure activities after having prepared this presentation with transparencies. Assessors were not allowed to question the assesseees. In the second presentation, assesseees had to prepare and conduct a presentation to convince a potential customer of the value of a product that they had received specific information about when

preparing for the exercise. The first leaderless group discussion was a hidden profile task (i.e., a task including asymmetrical distribution of information). In this exercise, assessees discussed a staffing task that required them to collaborate with each other to identify the most qualified applicant for a vacant position. To find the best solution, they had to share their information. In the second group discussion, participants were asked to rank ten university marketing strategies individually. Afterwards, they discussed these strategies with the goal of attaining a collective solution.

We employed a between-subjects design with a nontransparent and a transparent condition. Assessees and assessors were assigned to either the transparent condition or the nontransparent condition. Only the facilitators of the ACs and the first author of this study had access to this information. Since the assignment was based upon assessee's and assessors' availability, it was not random. No information was given to assesseees and assessors about the purpose of the study and the condition they had signed up for. Hence, we employed a quasi-experimental design with a double-blind procedure. To help make the two conditions comparable, we rotated the condition after every second weekend, such that after two weekends with transparent dimensions (i.e., transparent condition), over the next two weekends, dimensions were nontransparent (i.e., nontransparent condition).

Transparency was manipulated in the same way as in earlier studies by Kleinmann et al. (1996) and Smith-Jentsch et al. (2001). Hence, participants in the nontransparent condition received no information about the targeted dimensions. In contrast, participants in the transparent condition were informed before the preparation for the first AC exercise about the dimensions that were assessed in the AC. This was done in the preparation room by presenting a list of all dimensions that were measured in the AC and took the facilitator about 2 minutes. Thereafter, every time assesseees prepared for an exercise, they received an additional list with the targeted dimensions in the respective exercise from the facilitator who was monitoring participants' preparation. For each dimension, this list provided behavioral

examples for the dimension (e.g., Organizing & Planning: systematical and well-ordered approach, orders information according to criteria, initiates a plan of action/structure etc.), which were not tailored to the specific exercise (see Appendix B for an example of a transparency manipulation sheet). In both conditions, the preparation time for each AC exercise was exactly the same.

Measures

AC performance. For each exercise, two assessors rated participants' performance on a 5-point scale from 1 = *poor performance* to 5 = *excellent performance* on each dimension that was measured. After participants had completed all exercises, assessors discussed their ratings when they differed by two or more points and adjusted their ratings accordingly. Afterwards, we averaged the ratings of the two assessors. We computed overall dimension ratings (across all exercises per dimension) and an overall AC performance rating (across all exercises and all dimensions). To assess inter-rater reliability, we calculated intraclass correlations (ICC 1,1) between the overall dimension ratings of the two assessors for each exercise. The mean inter-rater reliability (i.e., the reliability of a single assessor) across all dimensions and all exercises was .72 before the discussion of the ratings and .82 after the discussion of the ratings.

Job performance and supervisors' demographic data. Around the time of the AC, participants' supervisors were asked to evaluate participants' job performance in an online questionnaire. Task performance was measured on a 7-point scale ranging from 1 = *not at all* to 7 = *absolutely*, with five items (see Appendix C for the items) from the task-based job performance questionnaire by Bott, Svyantek, Goodman, and Bernal (2003) translated by Jansen et al. (2013). Internal consistency (Cronbach's α) was .84 for this scale. OCB was measured on a 7-point scale ranging from 1 = *not at all* to 7 = *absolutely*, with fifteen items (see Appendix C for the items) from the German OCB scale by Staufenbiel and Hartz (2000). Internal consistency was .84 for this scale. Additionally, supervisors completed descriptive

items on their age, the number of years in a managing position, and how well they could evaluate the performance of their employee. Of the supervisors contacted, 90.9% answered the questionnaire.

Impression management. The self-promotion and ingratiation scales were based upon the taxonomy from McFarland et al. (2005). Three subject matter experts checked the items for their suitability in the AC exercise context. After participants completed all the AC exercises, we asked them to evaluate their use of self-promotion and ingratiation across the AC (i.e., with regard to all AC exercises, not separately for the respective AC exercises) on a scale from 1 = *I totally disagree* to 5 = *I fully agree* (see Appendix C for the items). Assesseees were informed that this measure was for research purposes only and did not affect the choice of the best participant of the day so that they had no motivation to distort their answers. Internal consistency was .83 for self-promotion and .73 for ingratiation.

Opportunity to perform. Opportunity to perform was measured with a German version of the Chance to Perform subscale from the Selection Procedural Justice Scale by Bauer et al. (2001) that was adapted by the authors of this article to the AC context (see Appendix C for the items). Participants answered on a scale from 1 = *I totally disagree* to 5 = *I fully agree*. Internal consistency was $\alpha = .76$.

Cognitive ability. We measured cognitive ability to test whether assesseees from both conditions were comparable. Cognitive ability was assessed with the Wiener Matrizen Test 2 (WMT-2, Formann, Waldherr, & Piswanger, 2011), a normed matrix test based upon Cattell's fluid intelligence model. This test is often used in German-speaking countries as it is less time-consuming (20-30 minutes) compared to others common cognitive ability tests. Test results from the WMT-2 have been shown to relate closely to results from the Standard Progressive Matrices and the Intelligenz-Struktur-Test (Amthauer, Brocke, Liepmann, & Beauducel, 2001), with correlations of .92 for the former and .85 for the latter. Internal

consistency of the WMT-2 scores was $\alpha = .76$ and the average raw score in our sample was 14.22 ($SD = 3.01$).

Manipulation check and demographic variables. Following Kleinmann et al. (1996) and Kolk et al. (2003), assessees from both conditions received a list of 11 dimensions after the last AC exercise. Besides the targeted 5 dimensions, the list included distracter dimensions that had not been evaluated during the AC (e.g., Creativity, Open-Mindedness). Participants were instructed to check the dimensions (either in the columns *yes* or *no*) that had been evaluated in the overall AC. For each answer, a value of 1 was assigned when they checked *yes* for a dimension that had actually been assessed during the AC or when they checked *no* for a dimension that had not been assessed in the AC. A value of 0 was assigned when the assessee checked *no* for a dimension that had actually been assessed or when they checked *yes* for a dimension that had actually not been assessed. We computed a sum score across all 11 dimensions to check whether the manipulation was successful. Following Jansen et al. (2013), participants additionally filled out items on how realistic they perceived the AC. Lastly, they reported their work experience and the number of hours they worked per week on one item and responded to sociodemographic items.

Procedure

A maximum of 12 participants took part in each AC for an entire day. In the beginning, facilitators gave an introductory session for participants during which assessors were also present. In this session, facilitators briefly introduced the participants to the selection process and the assessors. Participants were provided with a job advertisement that described the job of a management trainee position and were instructed to keep this position in mind during the AC. To increase their motivation to perform, participants were informed that the best two assessees on each AC day would receive the local equivalent of about 120 and 80 US dollars, respectively. When the introduction of the AC was completed, participants first completed a cognitive ability test and then took part in the different AC exercises (see

Appendix A for an overview of the exercises and dimensions). After the AC exercises, participants were then asked to fill out a questionnaire and were informed that their answers would only be used for research purposes and were not part of the simulated selection process. This questionnaire included questions concerning the use of self-promotion, a manipulation check for the transparency manipulation, several sociodemographic items and some questions about the perceived authenticity of the simulated AC. Finally, at the end of the AC, participants received extensive feedback on their AC performance from the assessors.

In the transparent condition, one participant was not able to complete all the measures as she was arriving late and leaving early so that her data on the cognitive ability measure, the opportunity to perform measure, and the self-promotion measure were missing. In the nontransparent condition, one participant did not complete the opportunity to perform measure. The task performance measure and OCB measure were not completed by 7 supervisors in the transparent condition and by 11 supervisors in the nontransparent condition. A few participants in the transparent condition and in the nontransparent conditions did not report their age but we were able to reconstruct the age from four participants online so that there were only two missing values for age in each condition. We treated these data as missing data and did not use imputation methods or exclude participants due to missing data.

Results

Preliminary Analysis and Descriptives

First, we checked the data on the perceived realism of the simulated AC. Assesseees reported that they behaved as if they applied for a job ($M = 1.79$, $SD = 0.62$, on a scale from 1 = *I fully agree* to 4 = *I disagree*) and that they were able to imagine being a real applicant ($M = 1.96$, $SD = 0.76$, on a scale from 1 = *I fully agree* to 4 = *I disagree*).

Second, we examined whether the manipulation of transparency was successful. Participants in the transparent conditions received a higher sum score on the manipulation

check ($M = 8.95$, $SD = 1.76$) than participants in the nontransparent conditions ($M = 6.74$, $SD = 1.39$), $t(173.24) = 9.51$, $p < .001$, $d = 1.39$. Thus, the manipulation was successful.

Third, given the argument that transparency might affect performance ratings, we examined whether AC performance ratings differed between conditions. Independent samples t -tests revealed no difference between the two conditions, $t(192) = 0.28$, $p = .779$, which is in line with the findings from the majority of studies (Kolk et al., 2003; Schulze Versmar et al., 2007; Smith-Jentsch, 2007, Study 2; Strobel et al., 2006).

Finally, we tested whether assessees in both conditions were comparable with regard to age, cognitive ability, sex, work experience, number of hours they worked per week, task performance ratings, OCB ratings, and how well supervisors reported to be able to evaluate their employees' performance. Independent samples t -tests revealed no differences between conditions for any of the variables, all $ts < 1.26$, and all $ps > .21$. Hence, assessees from both conditions were comparable.¹

Effects of Transparency on Criterion-related Validity

Hypothesis 1 stated that ratings from a nontransparent AC would be more criterion valid for predicting task performance on the job than ratings from a transparent AC. In line with this, the validity of overall AC performance ratings was significant in the nontransparent AC, $r = .24$, and nonsignificant in the transparent AC, $r = .08$. Moreover, four of the five correlations between AC dimension ratings and supervisor's task performance ratings were higher in the nontransparent AC than in the transparent AC (see Table 1). In line with this, a paired samples t -test revealed a significant difference between these five paired correlations (i.e. the five correlations in the nontransparent AC and the five correlations in the transparent

¹ Additionally, we also tested whether construct-related validity improved when dimensions were made transparent. Descriptively, using an r -to- Z transformation, the mean convergent validity across all coefficients was .33 in the transparent AC and .26 in the nontransparent AC and the mean discriminant validity across all coefficients was .52 in the transparent and .53 in the nontransparent AC. Using all convergent correlations as dependent variables and comparing them across the two conditions, we found that the difference was significant, $t(14) = 3.08$, $p = .008$. Concerning discriminant validity, the difference was not significant, $t(18) = 1.23$, $p = .237$. Hence, results from the present study replicate findings by Kleinmann et al. (1996) and Kolk et al. (2003, Study 2) that showed an improvement of convergent validity in the transparent AC, but no difference with regard to discriminant validity.

AC), $t(4) = 2.86$, $p = .045$, which provides support for Hypothesis 1. We further explored the dimension-level validities for the nontransparent and transparent AC. The pattern of dimension-level validities differed between the two conditions, for instance dimensions that had the second and third best validities in the nontransparent AC were then among the dimensions with the lowest validities in the transparent AC (organizing and planning, and persuasion). For instance, in agreement with meta-analytic results on the validity of dimensions (Arthur, Day, McNelly, & Edens, 2003), ratings of analytical skills were the best predictor of task performance in the nontransparent AC. In the transparent AC, presentation skills had the highest relation with job performance, but this relation was not significant, $p = .167$.

As a complementary analysis, we tested whether ratings from a nontransparent AC would be more criterion valid in predicting OCB. We first computed correlations between different AC dimension ratings and supervisors' OCB ratings (Table 1). The mean validity of overall AC performance ratings was not larger in the nontransparent than in the transparent AC, $r_s = .06$ and $.08$, respectively.

Effects of Transparency on the Relation between Impression Management and Job Performance

Hypothesis 2 predicted that transparency moderates the relation between self-promotion and job performance such that assessees' self-promotion in the nontransparent AC will be more positively related to supervisor-rated job performance than assessees' self-promotion in the transparent AC. Before testing this hypothesis, we examined whether there was a difference in the amount of self-promotion between the two conditions (i.e. the nontransparent and transparent AC). An independent samples t -test revealed no difference between the amount of self-promotion in the nontransparent ($M = 2.52$, $SD = 0.62$) versus in the transparent AC ($M = 2.57$, $SD = 0.77$), $t(190) = 0.53$, $p = .595$. This demonstrates that

participants in the transparent AC were not engaging in more self-promotion than those in the nontransparent AC.

Descriptively, the correlation of self-promotion and task performance was positive, $r = .24$, $p = .024$, for the nontransparent AC and negative, $r = -.20$, $p = .062$, for the transparent AC (Table 1). We conducted hierarchical multiple regressions to test Hypothesis 2. Transparency moderated the relation between self-promotion and task performance ($\Delta R^2 = .05$, $p = .003$, see Table 2) such that assessee's self-promotion in the nontransparent AC was more positively related to task performance on the job than assessee's self-promotion in the transparent AC (see Figure 1). These results provide support for Hypothesis 2.

Before testing the same effect for ingratiation, we also examined whether there was a difference between the two conditions. An independent samples t -test showed that the amount of ingratiation in the nontransparent ($M = 2.89$, $SD = 0.74$) was smaller than in the transparent AC ($M = 3.10$, $SD = 0.73$), $t(190) = 1.98$, $p = .049$. A hierarchical multiple regression analysis showed that transparency did not moderate the relation between ingratiation and task performance on the job ($\Delta R^2 = .00$, $p = .690$, see Table 2).

Effects of Transparency on the Opportunity to Perform

Hypothesis 3 stated that assessee's perceived opportunity to perform would be higher in the transparent AC than in the nontransparent AC. Descriptively, the average score was only minimally larger in the transparent ($M = 3.30$, $SD = 0.52$) than in the nontransparent AC ($M = 3.26$, $SD = 0.60$), and the independent samples t -test revealed no difference, $t(189) = 0.59$, $p = .558$. Hence, our results did not lend support to Hypothesis 3.

Discussion

The present study investigated the effects of making the targeted dimensions transparent to AC participants and thereby made several major contributions to the AC literature and the fairness perceptions literature. First of all, results relating to Hypothesis 1 illustrated that the criterion-related validity of a nontransparent AC that predicts task

performance decreased when dimensions were made transparent, thus providing an empirical answer to the debate on whether transparency has a negative or a positive impact on criterion-related validity. This study supports the conceptual assumptions drawn from CAPS theory that making dimensions transparent represents a deviation from the situational characteristics in the criterion and can hence lower criterion-related validity. The present results also provide empirical support for the speculation that transparency could reduce the criterion-related validity of selection procedures such as interviews and ACs (Klehe et al., 2008; Kleinmann, 1997; Smith-Jentsch, 2007).

Findings from the present study thus indicate that making dimensions transparent does not seem advisable in an AC for selection purposes, as revealing dimensions leads to AC ratings that do not predict supervisor-rated task performance as well as when dimensions are not revealed. When predicting OCB, however, it made no difference whether AC ratings stemmed from the transparent or nontransparent AC as neither were valid predictors. This might imply that ACs might better predict task-related performance than more discretionary types of behavior. Second, dimension transparency does not alter the relation to more discretionary types of criteria that capture behavior outside the task description.

A second implication of these findings relates to maintaining the non-transparency of dimensions in organizations that frequently conduct ACs with the same dimensions when selecting candidates. Given that organizations often provide feedback on assessee's performance on the targeted dimensions, this might increase the transparency of dimensions for future ACs. This has implications for internal ACs for promotion, and for ACs used for both internal and external candidates. For internal ACs for promotions, transparency might increase for all internal candidates over time. Still, we assume that the impact of knowledge about the targeted dimensions might depend on the extent to which assessee receive information about what is assessed in each exercise with concrete behavioral examples (as opposed to general information about what is assessed in the overall AC). For ACs with both

internal and external candidates, a larger problem might evolve given that the degree of transparency might be higher for internal than for external candidates (e.g., due to communication within the organization). We recommend that this issue be monitored by continuously evaluating the AC's criterion-related validity.

In addition to the primary finding related to the criterion-related validity, this study also illustrates that the impact of transparency on performance and validity in the AC is rather complex. This is because although we found detrimental effects of transparency on criterion-related validity, we replicated findings from the majority of previous studies that making dimensions transparent does not appear to increase mean AC scores (Kolk et al., 2003; Schulze Versmar et al., 2007; Smith-Jentsch, 2007, Study 2; Strobel et al., 2006). The lack of overall performance differences may be due to assessee's interindividual differences that interact with transparency, resulting in effects on AC performance that may vary in direction (positive effects for some assessee's, negative for others). We advocate studying the interplay of individual differences (e.g., test anxiety, perceived controllability of the situation), transparency, and AC performance to further enhance our understanding of the interactionist process.

This study also provides an answer to the issue of whether transparency causes assessee's to exhibit behaviors they otherwise would not exhibit. Our results showed that transparency was a moderator of the relation between self-promotion and job performance such that assessee's self-promotion in the nontransparent AC was more positively related to supervisor-rated job performance. Most interestingly, there was a negative relationship between self-promotion and job performance in the transparent AC. These findings hint at potentially detrimental effects of transparency given that self-promotion is positively related to AC performance, but negatively related to job performance. Hence, these results imply that transparency might encourage assessee's to engage in different types of self-promotion than they usually do on the job. In contrast, the positive relation between self-promotion and job

performance in the nontransparent AC suggests that self-promotion in the AC might be more typical for assessees (and hence less biased) and would also be used on the job. Vohs, Baumeister, and Ciarocco (2005) distinguish habitual or typical IM from IM that deviates from an individual's habits. Based upon the results of the present study, it seems useful to employ this distinction in the selection context and to take the typicality of the IM into consideration when examining effects of IM on job performance.

Related to this, findings from this study also indicate that situational features like transparency may impact different impression management behaviors in distinct ways. In contrast to self-promotion, transparency did not moderate the relation between ingratiation and job performance, and the level of ingratiation was higher in the transparent AC as compared to the nontransparent AC.

A final contribution of this study was the test of whether transparency affects assessees' perception of their opportunity to perform. From the assessees' perspective and based upon research on procedural fairness, it seemed plausible that informing assessees about targeted dimensions could influence assessees' perception of their opportunity to perform. However, our results did not support this assumption as there was no significant difference in assessees' perception of their opportunity to perform. One reason for this finding may be that ACs are characterized by a high level of opportunity to perform compared to some other types of selection procedures. As a consequence, it might be more difficult to further increase this perception by enhancing transparency.

Limitations

One issue that is important to address is that we collected data in a simulated setting and not in an operational AC. Participants in the AC held a wide variety of jobs and the AC was not designed to predict performance in any of their particular jobs. As such, relevance of dimensions and exercises might have varied across the jobs participants held. Nevertheless, our data showed that assessees perceived the setting as realistic. Further, manipulating

transparency in an applied setting could have led to internal validity problems as the double-blind design could have been less easily established and ethical reasons hindered us from conducting transparent ACs in the field given our assumptions that transparency lowers criterion-related validity.

A second issue relates to the non-random assignment of participants based upon their availability. To avoid systematic differences between participants of both conditions, we rotated the order of the transparent and nontransparent condition during the data collection and conducted the ACs for both conditions on the same days. Further, our data analysis revealed no difference between participants in the two conditions on any of the potentially relevant variables (e.g., age, gender, general mental ability, work experience, job performance ratings).

Another concern refers to manner in which IM was measured. We asked participants to report IM across the AC and not for each exercise separately to avoid triggering IM by asking about it during the AC. As such, our data did not capture potential variations of IM across exercises that would have allowed further analysis on the exercise level.

Future Research

The finding that ratings from a transparent AC were less criterion valid than ratings from a nontransparent AC illustrates that further research should continue examining what exactly explains this effect. Specifically, future research should focus on the difference between assessee's behavior in a transparent AC versus a nontransparent AC that might explain the different relation of AC performance to job performance. We provided a possible explanation that assumes that the relation of IM and job performance depends on how typical the IM shown is for a person. With regard to IM, we advocate further research that continues examining IM and the distinction between typical and situationally-based IM across contexts and sources. Accordingly, IM research from the selection context (e.g., Barrick et al., 2009; Klehe et al., 2014; McFarland et al., 2005) needs to be linked to IM research in the work

context (e.g., Higgins, Judge, & Ferris, 2003). Up to now, both research streams have been rather separate and we believe that shedding light on IM as behavior in different contexts will foster our understanding of how much variance in IM is situation-based and how much is person-based. Moreover, recent research on IM in the interview has indicated the need to take different sources (e.g., interviewers, interviewees) into account when studying IM (e.g., Ingold, Kleinmann, König, & Melchers, 2014; Roulin, Bangerter, & Levashina, 2014). Hence, continuing research should explore differences in IM as perceived by different sources (e.g., external observers in the AC, supervisors, and peers in the work context).

On a related note, future research should examine whether the degree of dimension transparency relates to criterion-related validity. Specifically, we predict that when information about the targeted dimensions is less salient than in the present study (i.e., when the dimensions are only mentioned in the beginning), correlations between AC performance and job performance will be greater in magnitude. Although surveys have indicated that different degrees of transparency are employed in ACs and that the percentage of ACs that inform assesseees about targeted dimension has increased (Thornton & Krause, 2009), these have not been considered in primary or meta-analytic research on AC criterion-related validity. For this reason, we recommend further primary research that systematically manipulates and reports the degree of transparency and examines its effect on criterion-related validity as well as meta-analytic research that considers the effects of transparency in operational settings.

We recommend extending research on transparency by examining how providing information about exercises and not only dimensions affects criterion-related validity. This relates to the discussion on whether effective AC performance should be conceptualized with regard to dimensions, exercises, or some combination of the two (e.g., Lance, 2008; Lance, Lambert, Gewin, Lievens, & Conway, 2004). Starting this line of research may provide more

insights into the mechanisms underlying the criterion-related validity of ACs, as well as inform the dimension versus exercise debate.

References

- Amthauer, R., Brocke, B., Liepmann, D., & Beauducel, A. (2001). *Intelligenz-Struktur-Test 2000 R [Intelligence Structure Test 2000 R]*. Göttingen, Germany: Hogrefe.
- Anderson, N. (1991). Decision making in the graduate selection interview: An experimental investigation. *Human Relations, 44*, 403-417. doi:10.1177/001872679104400407
- Anderson, N., & Goltsi, V. (2006). Negative psychological effects of selection methods: Construct formulation and an empirical investigation into an assessment center. *International Journal of Selection and Assessment, 14*, 236-255. doi:10.1111/j.1468-2389.2006.00344.x
- Arthur, W., Day, E. A., McNelly, T. L., & Edens, P. S. (2003). A meta-analysis of the criterion-related validity of assessment center dimensions. *Personnel Psychology, 56*, 125-154. doi:10.1111/j.1744-6570.2003.tb00146.x
- Barrick, M. R., Shaffer, J. A., & DeGrassi, S. W. (2009). What you see may not be what you get: Relationships among self-presentation tactics and ratings of interview and job performance. *Journal of Applied Psychology, 94*, 1394-1411. doi:10.1037/a0016532
- Bauer, T. N., Truxillo, D. M., Sanchez, R. J., Craig, J. M., Ferrara, P., & Campion, M. A. (2001). Applicant reactions to selection: Development of the selection procedural justice scale (SPJS). *Personnel Psychology, 54*, 387-419. doi:10.1111/j.1744-6570.2001.tb00097.x
- Bott, J. P., Svyantek, D. J., Goodman, S. A., & Bernal, D. S. (2003). Expanding the performance domain: Who says nice guys finish last? *International Journal of Organizational Analysis, 11*, 137-152. doi:10.1108/eb028967
- Dodd, W. E. (1977). Attitudes toward assessment center programs. In J. I. Moses & W. C. Byham (Eds.), *Applying the assessment center method* (pp. 1631-1183). New York: Pergamon.

- Ellis, A. P., West, B. J., Ryan, A. M., & DeShon, R. P. (2002). The use of impression management tactics in structured interviews: A function of question type? *Journal of Applied Psychology, 87*, 1200-1208. doi:10.1037/0021-9010.87.6.1200
- Eurich, T. L., Krause, D. E., Cigularov, K., & Thornton, G. C. (2009). Assessment centers: Current practices in the United States. *Journal of Business and Psychology, 24*, 387-407. doi:10.1007/s10869-009-9123-3
- Formann, A. K., Waldherr, K., & Piswanger, K. (2011). *Wiener Matrizen-Test 2 [Vienna Matrices Test 2]*. Göttingen, Germany: Beltz.
- Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review, 18*, 694-734.
doi:10.5465/amr.1993.9402210155
- Griffin, B. (2014). The ability to identify criteria: Its relationship with social understanding, preparation, and impression management in affecting predictor performance in a high-stakes selection context. *Human Performance, 27*, 147-164.
doi:10.1080/08959285.2014.882927
- Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology, 57*, 639-683.
- Higgins, C. A., Judge, T. A., & Ferris, G. R. (2003). Influence tactics and work outcomes: A meta-analysis. *Journal of Organizational Behavior, 24*, 89-106. doi:10.1002/job.181
- Hoffman, B. J. (2014, October). *Assessment centers are deficient predictors of effective leadership*. Paper presented at the 38th International Congress on Assessment Center Methods, Alexandria, US.
- Ingold, P. V., Kleinmann, M., König, C. J., & Melchers, K. G. (2014). Shall we continue or stop disapproving of self-presentation? Evidence on impression management and faking in a selection context and their relation to job performance. *European Journal*

of Work and Organizational Psychology. Advance online publication.

doi:10.1080/1359432X.2014.915215

Ingold, P. V., Kleinmann, M., König, C. J., Melchers, K. G., & Van Iddekinge, C. H. (2014).

Why do situational interviews predict job performance: The role of interviewees' ability to identify criteria. *Journal of Business and Psychology*. Advance online publication. doi:10.1007/s10869-014-9368-3

International Task Force on Assessment Center Guidelines. (2009). Guidelines and ethical considerations for assessment center operations. *International Journal of Selection and Assessment*, 17, 243-253. doi:10.1111/j.1468-2389.2009.00467.x

Jansen, A., Lievens, F., & Kleinmann, M. (2011). Do individual differences in perceiving situational demands moderate the relationship between personality and assessment center dimension ratings? *Human Performance*, 24, 231-250.

doi:10.1080/08959285.2011.580805

Jansen, A., Melchers, K. G., Lievens, F., Kleinmann, M., Brändli, M., Fraefel, L., et al.

(2013). Situation assessment as an ignored factor in the behavioral consistency paradigm underlying the validity of personnel selection procedures. *Journal of Applied Psychology*, 98, 326-341. doi:10.1037/a0031257

Klehe, U.-C., Kleinmann, M., Niess, C., & Grazi, J. (2014). Impression management behavior during assessment centers: Artificial behavior or much ado about nothing? *Human Performance*, 27, 1-24. doi:10.1080/08959285.2013.854365

Klehe, U.-C., König, C. J., Richter, G. M., Kleinmann, M., & Melchers, K. G. (2008).

Transparency in structured interviews: Consequences for construct and criterion-related validity. *Human Performance*, 21, 107-137. doi:10.1080/08959280801917636

Kleinmann, M. (1993). Are rating dimensions in assessment centers transparent for participants? Consequences for criterion and construct validity. *Journal of Applied Psychology*, 78, 988-993. doi:10.1037/0021-9010.78.6.988

- Kleinmann, M. (1997). Transparenz der Anforderungsdimensionen: Ein Moderator der Konstrukt- und Kriteriumsvalidität [Transparency of the required dimensions: A moderator of assessment centers' construct and criterion validity]. *Zeitschrift für Arbeits- und Organisationspsychologie*, *41*, 171-181.
- Kleinmann, M., Ingold, P. V., Lievens, F., Jansen, A., Melchers, K. G., & König, C. J. (2011). A different look at why selection procedures work: The role of candidates' ability to identify criteria. *Organizational Psychology Review*, *1*, 128-146.
doi:10.1177/2041386610387000
- Kleinmann, M., & Klehe, U.-C. (2011). Selling oneself: Construct and criterion-related validity of impression management in structured interviews. *Human Performance*, *24*, 29-46. doi:10.1080/08959285.2010.530634
- Kleinmann, M., Kuptsch, C., & Köller, O. (1996). Transparency: A necessary requirement for the construct validity of assessment centres. *Applied Psychology: An International Review*, *45*, 67-84. doi:10.1111/j.1464-0597.1996.tb00849.x
- Kolk, N. J., Born, M. P., & van der Flier, H. (2003). The transparent assessment centre: The effects of revealing dimensions to candidates. *Applied Psychology: An International Review*, *52*, 648-668. doi:10.1111/1464-0597.00156
- König, C. J., Klehe, U.-C., Berchtold, M., & Kleinmann, M. (2010). Reasons for being selective when choosing personnel selection procedures. *International Journal of Selection and Assessment*, *18*, 17-27. doi:10.1111/j.1468-2389.2010.00485.x
- König, C. J., Melchers, K. G., Kleinmann, M., Richter, G. M., & Klehe, U.-C. (2007). Candidates' ability to identify criteria in nontransparent selection procedures: Evidence from an assessment center and a structured interview. *International Journal of Selection and Assessment*, *15*, 283-292. doi:10.1111/j.1468-2389.2007.00388.x
- Krause, D. E., & Gebert, D. (2003). A comparison of assessment center practices in organizations in German-speaking regions and the United States. *International*

- Journal of Selection and Assessment*, 11, 297-312. doi:10.1111/j.0965-075X.2003.00253.x
- Krause, D. E., Rossberger, R. J., Dowdeswell, K., Venter, N., & Joubert, T. (2011). Assessment center practices in South Africa. *International Journal of Selection and Assessment*, 19, 262-275. doi:10.1111/j.1468-2389.2011.00555.x
- Krause, D. E., & Thornton, G. C. (2009). A cross-cultural look at assessment center practices: Survey results from western europe and north america. *Applied Psychology: An International Review*, 58, 557-585. doi:10.1111/j.1464-0597.2008.00371.x
- Kristof-Brown, A., Barrick, M. R., & Franke, M. (2002). Applicant impression management: Dispositional influences and consequences for recruiter perceptions of fit and similarity. *Journal of Management*, 28, 27-46. doi:10.1177/014920630202800103
- Kurecka, P. M., Austin, J. M., Johnson, W., & Mendoza, J. L. (1982). Full and errant coaching effects on assigned role leaderless group discussion performance. *Personnel Psychology*, 35, 805-812. doi:10.1111/j.1744-6570.1982.tb02223.x
- Lance, C. E. (2008). Why assessment centers do not work the way they are supposed to. *Industrial and Organizational Psychology*, 1, 84-97. doi:10.1111/j.1754-9434.2007.00017.x
- Lance, C. E., Lambert, T. A., Gewin, A. G., Lievens, F., & Conway, J. M. (2004). Revised estimates of dimension and exercise variance components in assessment center postexercise dimension ratings. *Journal of Applied Psychology*, 89, 377-385. doi:10.1037/0021-9010.89.2.377
- Levashina, J., & Campion, M. A. (2007). Measuring faking in the employment interview: Development and validation of an interview faking behavior scale. *Journal of Applied Psychology*, 92, 1638-1656. doi:10.1037/0021-9010.92.6.1638

- Levashina, J., Hartwell, C. J., Morgeson, F. P., & Campion, M. A. (2014). The structured employment interview: Narrative and quantitative review of the research literature. *Personnel Psychology, 67*, 241–293. doi:10.1111/peps.12052
- Lievens, F., Harris, M. M., Van Keer, E., & Bisqueret, C. (2003). Predicting cross-cultural training performance: The validity of personality, cognitive ability, and dimensions measured by an assessment center and a behavior description interview. *Journal of Applied Psychology, 88*, 476-489. doi:10.1037/0021-9010.88.3.476
- Lievens, F., Tett, R. P., & Schleicher, D. J. (2009). Assessment centers at the crossroads: Toward a reconceptualization of assessment center exercises. In J. L. Martocchio & H. Liao (Eds.), *Research in Personnel and Human Resource Management* (pp. 99-152). Bingley, UK: Emerald.
- Maurer, T. J., Solamon, J. M., & Lippstreu, M. (2008). How does coaching interviewees affect the validity of a structured interview? *Journal of Organizational Behavior, 29*, 355-371. doi:10.1002/Job.512
- McFarland, L. A., Ryan, A. M., & Kriska, S. D. (2002). Field study investigation of applicant use of influence tactics in a selection interview. *The Journal of Psychology, 136*, 383–398. doi:10.1080/00223980209604165
- McFarland, L. A., Ryan, A. M., & Kriska, S. D. (2003). Impression management use and effectiveness across assessment methods. *Journal of Management, 29*, 641-661. doi:10.1016/S0149-2063(03)00030-8
- McFarland, L. A., Yun, G., Harold, C. M., Viera, L., & Moore, L. G. (2005). An examination of impression management use and effectiveness across assessment center exercises: The role of competency demands. *Personnel Psychology, 58*, 949-980. doi:10.1111/j.1744-6570.2005.00374.x
- Melchers, K. G., Ingold, P. V., Wilhelmy, A., & Kleinmann, M. (in press). Beyond validity: Shedding light on the social situation in employment interviews. In I. Nikolaou & J.

- K. Oostrom (Eds.), *Employee Recruitment, Selection, and Assessment: Contemporary Issues for Theory and Practice*. Hove, UK: Psychology Press.
- Melchers, K. G., Klehe, U.-C., Richter, G. M., Kleinmann, M., König, C. J., & Lievens, F. (2009). "I know what you want to know": The impact of interviewees' ability to identify criteria on interview performance and construct-related validity. *Human Performance, 22*, 355-374. doi:10.1080/08959280903120295
- Melchers, K. G., Wirz, A., & Kleinmann, M. (2012). Dimensions AND exercises: Theoretical background of mixed-model assessment centers. In D. J. R. Jackson, C. E. Lance & B. J. Hoffman (Eds.), *The psychology of assessment centers* (pp. 237-254). New York: Routledge.
- Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*, 246-268. doi:10.1037/0033-295X.102.2.246
- Organ, D. W. (1988). A restatement of the satisfaction-performance hypothesis. *Journal of Management, 14*, 547-557. doi:10.1177/014920638801400405
- Preckel, D., & Schüpbach, H. (2005). Zusammenhänge zwischen rezeptiver Selbstdarstellungskompetenz und Leistung im Assessment Center [Correlations between receptive self-presentation competence and performance in an assessment center]. *Zeitschrift für Personalpsychologie, 4*, 151-158. doi:10.1026/1617-6391.4.4.151
- Roch, S. G., Woehr, D. J., Mishra, V., & Kieszczyńska, U. (2012). Rater training revisited: An updated meta-analytic review of frame-of-reference training. *Journal of Occupational and Organizational Psychology, 85*, 370-395. doi:10.1111/j.2044-8325.2011.02045.x

- Roulin, N., Bangerter, A., & Levashina, J. (2014). Interviewers' perceptions of impression management in employment interviews. *Journal of Managerial Psychology, 29*, 141-163. doi:10.1108/JMP-10-2012-0295
- Schleicher, D. J., Venkataramani, V., Morgeson, F. P., & Campion, M. A. (2006). So you didn't get the job ... now what do you think? Examining opportunity-to-perform fairness perceptions. *Personnel Psychology, 59*, 559-590. doi:10.1111/j.1744-6570.2006.00047.x
- Schlenker, B. R. (1980). *Impression management*. Monterey, CA: Brooks/Cole.
- Schuler, H., Hell, B., Trapmann, S., Schaar, H., & Boramir, I. (2007). Die Nutzung psychologischer Verfahren der externen Personalauswahl in deutschen Unternehmen. Ein Vergleich über 20 Jahre [Use of personnel selection instruments in German organizations in the last 20 years]. *Zeitschrift für Personalpsychologie, 6*, 60-70. doi:10.1026/1617-6391.6.2.60
- Schulze Versmar, K., Thomas, E., & Kersting, M. (2007). Wissen, was gefordert wird [Knowing the demands]. *PersonalMagazin, 8*, 20-22.
- Smith-Jentsch, K. A. (2007). The impact of making targeted dimensions transparent on relations with typical performance predictors. *Human Performance, 20*, 187-203. doi:10.1080/08959280701332992
- Smith-Jentsch, K. A., Salas, E., & Brannick, M. T. (2001). To transfer or not to transfer? Investigating the combined effects of trainee characteristics, team leader support, and team climate. *Journal of Applied Psychology, 86*, 279-292. doi:10.1037/0021-9010.86.2.279
- Speer, A. B., Christiansen, N. D., König, C. J., Melchers, K. G., & Kleinmann, M. (2014). Establishing the cross-situational convergence of the ability to identify criteria: Consistency and prediction across similar and dissimilar assessment center exercises. *Human Performance, 27*, 44-60. doi:10.1080/08959285.2013.854364

- Spychalski, A. C., Quinones, M. A., Gaugler, B. B., & Pohley, K. (1997). A survey of assessment center practices in organizations in the United States. *Personnel Psychology, 50*, 71-90. doi:10.1111/j.1744-6570.1997.tb00901.x
- Staufenbiel, T., & Hartz, C. (2000). Organizational Citizenship Behavior: Entwicklung und erste Validierung eines Meßinstruments [Organizational citizenship behavior: Development and validation of a measurement instrument]. *Diagnostica, 46*, 73-83. doi:10.1026//0012-1924.46.2.73
- Stevens, C., & Kristof, A. (1995). Making the right impression: A field study of applicant impression management during job interviews. *Journal of Applied Psychology, 80*, 587-606. doi:10.1037/0021-9010.80.5.587
- Strobel, A., Lerche, S., Hoppe, K., & Bolte, E.-A. (2006). "Transparenz im AC - Aussagekraft ade?" Ein Beitrag zur Transparenzdiskussion im Assessment Center [Transparency in assessment centers - A farewell to quality? An essay about transparency and its consequences in assessment centers]. *Wirtschaftspsychologie, 8*, 48-57.
- Thornton, G. C., & Gibbons, A. M. (2009). Validity of assessment centers for personnel selection. *Human Resource Management Review, 19*, 169-187. doi:10.1016/j.hrmr.2009.02.002
- Thornton, G. C., & Krause, D. E. (2009). Selection versus development assessment centers: An international survey of design, execution, and evaluation. *International Journal of Human Resource Management, 20*, 478. doi:10.1080/09585190802673536
- Truxillo, D. M., & Bauer, T. N. (2011). Applicant reactions to organizations and selection systems. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology* (pp. 379-397). Washington, DC: American Psychological Association.
- Vohs, K. D., Baumeister, R. F., & Ciarocco, N. J. (2005). Self-regulation and self-presentation: Regulatory resource depletion impairs impression management and

- effortful self-presentation depletes regulatory resources. *Journal of Personality and Social Psychology*, 88, 632-657. doi:10.1037/0022-3514.88.4.632
- Wayne, S. J., & Kacmar, K. M. (1991). The effects of impression management on the performance appraisal process. *Organizational Behavior and Human Decision Processes*, 48, 70-88. doi:10.1016/0749-5978(91)90006-F
- Wayne, S. J., & Liden, R. C. (1995). Effects on impression management on performance ratings: A longitudinal study. *Academy of Management Journal*, 38, 232-260. doi:10.2307/256734
- Wernimont, P. F., & Campbell, D. T. (1968). Signs, samples and criteria. *Journal of Applied Psychology*, 52, 372-376. doi:10.1037/h0026244
- Wirz, A., Melchers, K., Schultheiss, S., & Kleinmann, M. (2014). Are improvements in assessment center construct-related validity paralleled by improvements in criterion-related validity: The effects of exercise similarity on assessment center validity. *Journal of Personnel Psychology*, 13, 184-193. doi:10.1027/1866-5888/a000115
- Woehr, D. J., & Huffcutt, A. I. (1994). Rater training for performance-appraisal: A quantitative review. *Journal of Occupational and Organizational Psychology*, 67, 189-205. doi:10.1111/j.2044-8325.1994.tb00562.x

Table 1
Descriptive Statistics and Intercorrelations of Study Variables in the Nontransparent and Transparent Condition

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	M	SD
1. Task-based job performance		.66**	.08	.04	.09	-.01	.15	.06	.15	-.03	.11	.03	-.20 [†]	-.05	-.07	.09	-.13	.00	5.80	0.87
2. Organizational Citizenship Behavior	.49**		.03	.01	.01	.03	.03	.02	.06	-.09	.14	.03	-.11	.02	-.05	-.07	-.19 [†]	.08	5.83	0.62
3. Overall AC performance	.24*	.06		.70**	.57**	.89**	.79**	.88**	.75**	.77**	.74**	.74**	.26*	.25*	.16	.08	-.15	-.15	3.56	0.57
4. Analytical skills	.29*	.13	.68**		.23*	.62**	.54**	.55**	.40**	.74**	.65**	.31**	.15	.22*	.07	.23*	-.08	-.18 [†]	3.55	0.69
5. Consideration of others	.17	-.03	.61**	.39**		.35**	.37**	.43**	.33**	.26*	.54**	.64**	.14	.06	.07	.05	.07	-.14	3.61	0.66
6. Persuasion	.18 [†]	.01	.89**	.61**	.47**		.65**	.68**	.63**	.72**	.64**	.68**	.29**	.21*	.23*	.05	-.15	-.16	3.56	0.69
7. Presentation skills	.14	-.04	.67**	.30**	.32**	.47**		.61**	.66**	.74**	.49**	.48**	.05	.20 [†]	.12	.08	-.12	-.05	3.73	0.76
8. Organizing & planning	.19 [†]	.08	.91**	.54**	.43**	.74**	.57**		.68**	.69**	.70**	.61**	.25*	.25*	.10	.05	-.20 [†]	-.15	3.43	0.73
9. Presentation 1	.11	.09	.65**	.23*	.21*	.53**	.66**	.62*		.44**	.34**	.47**	.14	.20 [†]	.12	-.04	-.19 [†]	.05	3.76	0.78
10. Presentation 2	.11	-.02	.74**	.66**	.27*	.66**	.66**	.68**	.37**		.48**	.36**	.18 [†]	.17	.03	.20 [†]	-.08	-.18 [†]	3.62	0.83
11. Group Discussion 1	.25*	.04	.69**	.70**	.66**	.64**	.25*	.59**	.16	.36**		.43**	.08	.20 [†]	.19 [†]	.07	-.14	-.24*	3.38	0.72
12. Group Discussion 2	.20 [†]	.05	.72**	.31**	.58**	.68**	.31**	.66**	.30**	.34**	.41**		.37**	.18 [†]	.16	.03	-.03	-.08	3.50	0.72
13. Self-promotion	.24*	.20 [†]	.00	-.04	-.10	.03	-.36	.04	.04	-.06	.01	.00		.32**	.07	-.02	-.15	-.08	2.57	0.77
14. Ingratiation	.01	.00	.23*	.13	.24*	.21*	.24*	.15	.21*	.12	.21*	.11	.25*		.33**	-.13	-.17	.05	3.10	0.73
15. Opportunity to perform	.09	-.04	.17	.10	.20*	.16	.04	.12	.13	-.01	.17	.18 [†]	.03	.03		-.06	-.01	-.08	3.47	0.69
16. Cognitive ability	.16	-.02	.06	.11	.10	.06	.01	.02	-.18 [†]	.06	.11	.19 [†]	.08	.19 [†]	-.05		-.09	-.07	97.43	13.67
17. Age	.11	.03	-.01	.02	-.06	.05	.06	-.07	.04	.05	.03	-.17	-.01	-.26*	-.15	-.29**		.07	28.37	5.21
18. Sex	-.07	.00	-.24*	-.09	-.16	-.16	-.12	-.31**	-.15	-.19 [†]	-.14	-.21*	-.04	-.20 [†]	-.04	.22*	.24*		1.51	0.50
M	5.73	5.87	3.59	3.61	3.74	3.59	3.65	3.42	3.69	3.62	3.44	3.61	2.52	2.89	3.35	99.47	28.70	1.42		
SD	0.87	0.65	0.55	0.75	0.64	0.64	0.70	0.73	0.79	0.79	0.79	0.76	0.62	0.74	0.73	13.03	5.64	0.50		

Note. Intercorrelations for the nontransparent condition are presented below the diagonal, and intercorrelations for the transparent condition are presented above the diagonal. For the nontransparent condition $N = 100$ with the exception of job performance for which $N = 89$ and for age for which $N = 98$. For the transparent condition $N = 94$ with the exception of job performance for which $N = 87$ and for age for which $N = 92$. Sex was coded as 1 = male, 2 = female. AC = Assessment center. [†] $p < .10$, * $p < .05$, ** $p < .01$, two-tailed.

Table 2
Results of Hierarchical Regression Analyses of Transparency as Moderator of Relations between Impression Management and Job Performance

Model 1	Supervisor-Rated Task Performance		
	β	R^2	ΔR^2
Step 1:		.00	.00
Self-promotion	.00		
Transparency	-.03		
Step 2:		.05*	.05**
Interaction	.29**		
Model 2	β	R^2	ΔR^2
Step 1:		.00	.00
Ingratiation	-.02		
Transparency	-.04		
Step 2:		.00	.00
Interaction	.13		

Note. $N = 174$;

* $p < .05$, ** $p < .01$ two-tailed.

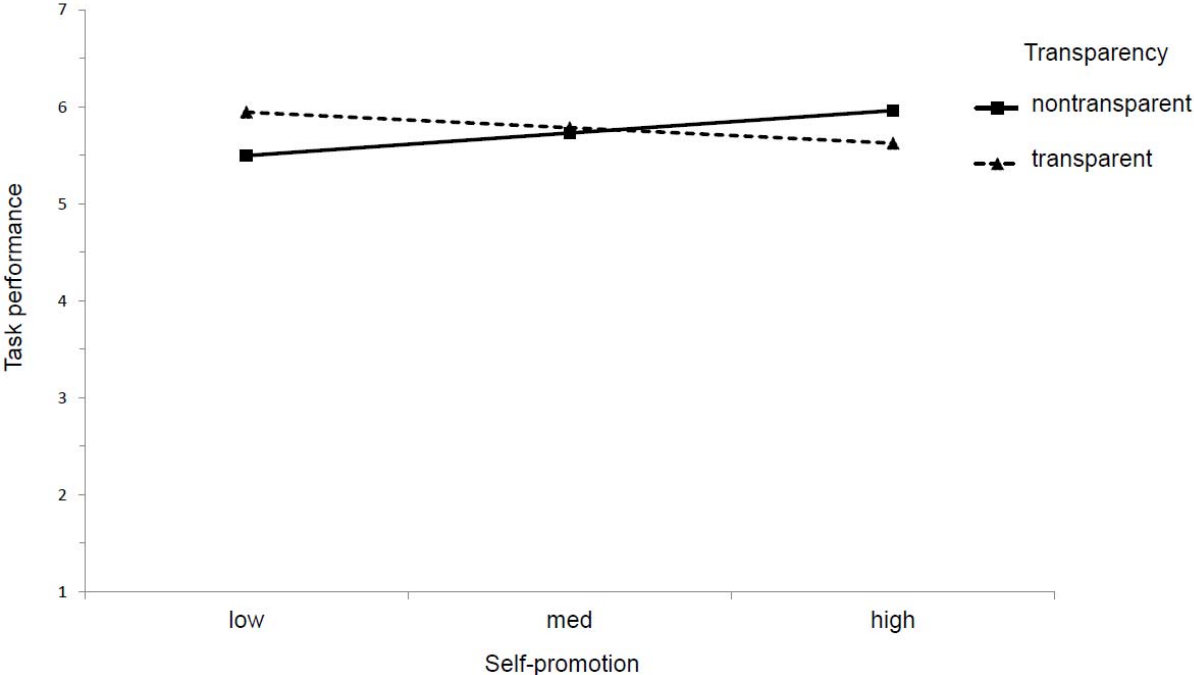


Figure 1. Interaction effects of transparency and assessesees' self-promotion in predicting supervisor-rated task performance.

Appendix A**Assessment Center Exercise-Dimension Matrix**

Table A.

Dimension by Exercise Matrix from the Assessment Center

Dimensions	Exercise			
	Presentation 1	Presentation 2	Group	Group
			discussion 1	discussion 2
Analytical skills		x	x	
Organizing and planning	x	x	x	x
Persuasiveness	x	x	x	x
Consideration of others			x	x
Presentation skills	x	x		

Appendix B

Transparency Manipulation Sheet for Group Discussion 2

Dimensions measured

1) Organizing and planning

Behavioral examples:

- proceed systematically and in a structured manner
- order information in a nuanced and comprehensive manner
- establish a clear, simple and adequate structure
- follow a common thread
- subdivide discussions and presentations
- correct themselves and others when deviating from the agreed procedure
- comply with time requirements

2) Persuasion

Behavioral examples:

- put forth vivid, concrete and nuanced arguments
- argue coherently
- give clear and understandable reasons for decisions
- argue in a consistent, committed and persistent manner
- verbally convince others
- sell own ideas

3) Consideration of others

Behavioral examples:

- pick up on opinions and ideas that contradict one's own opinion
- show willingness to depart from one's own position
- find common ground and emphasize unifying aspects
- put back own interests in favour of the common objectives
- actively work towards the joint achievement of goals
- endorse group decisions and align oneself with the group consensus

Appendix C

Measures of Job Performance, Impression Management, and Opportunity to Perform

1. Job performance measures

a) Task performance

1. The employee demonstrates expertise in all job-related tasks.
2. The employee manages more responsibility than typically assigned.
3. The employee fulfills all the requirements of the job.
4. The employee achieves the objectives of the job.
5. The employee plans and organizes to achieve objectives of the job.

b) Organizational citizenship behavior (OCB)

1. The employee helps others when they are overwhelmed by work.
2. The employee always comes to work on time.
3. The employee spends a lot of time complaining about trivial things.
4. The employee acts as a calming influence on colleagues when arguments arise.
5. The employee will inform well in advance if he or she is unable to come to work.
6. The employee tends to make a mountain out of a molehill.
7. The employee will voluntarily take the initiative to help new colleagues with on-the-job training.
8. The employee stands out due to exceptionally few absences from work.
9. The employee perceives everything the company does to be wrong.
10. The employee actively tries to prevent difficulties with colleagues.
11. The employee follows rules and work instructions to the letter.
12. The employee often criticizes colleagues.
13. The employee cheers up colleagues when they are feeling low.
14. The employee only takes time off for absolutely essential reasons.
15. The employee expresses reservations towards any changes within the company.

2. Impression management measures

a) Self-promotion

1. I have stated that I work very diligently.
2. I have stated that I have successfully dealt with similar situations.
3. I have emphasized my own strengths.
4. I have pointed out my own skills to interaction partners.
5. I have presented my own successes.
6. I have stated that I was largely responsible for successes.
7. I have stated that I contributed significantly to the success of successful projects.
8. I have stated that my proposed solution was already successful in a similar situation.
9. I have told interaction partners about problems that I had to solve in order to achieve a particular goal.

b) Ingratiation

1. I have praised an interaction partner.
2. I have given an interaction partner compliments.
3. I have praised the idea of an interaction partner.
4. I have praised the behavior of an interaction partner.

3. Opportunity to perform measure

1. I could really show my skills and abilities in these assessment center exercises.
2. These assessment center exercises give assessees the opportunity to show what they really can do.
3. These assessment center exercises offer applicants the possibility to show what they are really capable of.
4. It is possibly to show what one is capable of during these assessment center exercises.