Singapore Management University
Institutional Knowledge at Singapore Management University

Research Collection Lee Kong Chian School Of Business

Lee Kong Chian School of Business

9-2012

Responding to personality tests in a selection context: The role of the ability to identify criteria and the ideal-employee factor

Ute-Christine KELHE
Universitat Giessen

Martin KLEINMANN
University of Zurich

Thomas HARTSTEIN
Deka-Bank, Germany

Klaus G. MELCHERS
University of Zurich

Cornelius J. KONIG
Universitat des Saarlandes

See next page for additional authors

Follow this and additional works at: https://ink.library.smu.edu.sg/lkcsb_research

DOI: https://doi.org/10.1080/08959285.2012.703733

Part of the Human Resources Management Commons, and the Organizational Behavior and Theory Commons

Citation

Author

Ute-Christine KELHE, Martin KLEINMANN, Thomas HARTSTEIN, Klaus G. MELCHERS, Cornelius J. KONIG, Peter A. HESLIN, and Filip LIEVENS

Running head: Ideal-employee factor

Responding to Personality Tests in a Selection Context:

The Role of the Ability to Identify Criteria and the Ideal-Employee Factor

Abstract

Personality assessments are often distorted during personnel selection, resulting in a common "ideal-employee factor" (IEF) underlying ratings of theoretically unrelated constructs. This seems not to affect the personality measures' criterion-related validity, however. The current study attempts to explain this phenomenon by combining the literature on response distortion with the ones on cognitive schemata and on candidates' ability to identify criteria (ATIC). During a simulated selection process, 149 participants filled out Big Five personality measures and participated in several high- and low-fidelity work simulations to estimate their managerial performance. Structural equation modeling showed that the IEF presents an indicator of response distortion and that ATIC accounted for variance between the IEF and performance during the work simulations, even after controlling for self-monitoring and general mental ability.

Responding to Personality Tests in a Selection Context:

The Role of the Ability to Identify Criteria and the Ideal-Employee Factor

One of the definitions of faking is "saying what you think you ought to say rather

than what you really want to say." We have a word for that – "civilization."

(Murphy, in Morgeson et al., 2007a, p. 712)

Measures of the Big Five personality dimensions (Conscientiousness, Emotional Stability,

Openness to Experience, Extraversion, Agreeableness) can predict performance for many tasks

and jobs (Barrick, Mount, & Judge, 2001; Ones, Dilchert, Viswesvaran, & Judge, 2007; Tett &

Christiansen, 2007; though see Morgeson et al., 2007a, 2007b). Yet, during personnel selection,

these measures are vulnerable in applicants distorting responses (Morgeson et al., 2007a, 2007b;

Rosse, Stecher, Miller, & Levin, 1998; Whyte, 1956). Response distortion, also termed

impression management, socially desirable responding, faking good, or self-enhancement

(McFarland & Ryan, 2000; Morgeson et al., 2007b), happens when respondents "manipulate

responses to personality items to make a positive impression" (Zickar & Robie, 1999, p. 551).

Most selection literature concurs that response distortion happens during personnel

selection (Griffith & Peterson, 2006; Morgeson et al., 2007b; Tett & Christiansen, 2007;

Viswesvaran & Ones, 1999), yet agrees far less about whether such distortion reflects a social

skill (Murphy in Morgeson et al., 2007b; Rosse et al., 1998; Viswesvaran & Ones, 1999),

something inconsequential (Ones et al., 2007; Ones, Viswesvaran, & Reiss, 1996), or a shameful

scam (Campion in Morgeson et al., 2007b; Tett & Christiansen, 2007; see also Kuncel &

Borneman, 2007). What we do know (e.g., J. Hogan, Barrett, & Hogan, 2007) is that, first, the

effects of distortion usually differ across dimensions: A meta-analysis by Birkeland et al. (2006)

showed that real-life distortion is primarily linked to increased scores on Conscientiousness and

Emotional Stability, while score-increases on the other three major personality dimensions are

much smaller (this stands in contrast to laboratory studies where 'faking good' instructions often lead to considerable effects across all five dimensions; Viswesvaran & Ones, 1999). Second, response distortion affects the rank ordering of candidates during a top down selection (Mueller-Hanson, Heggestad, & Thornton, 2003). Third, however, response distortion seems unrelated to the personality tests' criterion-related validity in field settings (see Schmitt & Oswald, 2006). Fourth, response distortion impairs the personality scores' construct-related validity: In applicant samples, personality items load not only on their respective Big Five factor but also on a sixth factor, the "ideal-employee factor" (Schmit & Ryan, 1993). Although replicated in several applicant samples of personality ratings (Collins & Gleaves, 1998; Ellingson, Sackett, & Hough, 1999; Pauls & Crost, 2005; M. A. Smith, Moriatry, Lutrick, & Canger, 2001; Topping & O'Gorman, 1997), Cellar, Miller, Doverspike, and Klawsky (1996, p. 703) noted: "There is a sixth factor, but it is still not clear what the sixth factor is". One and a half decades later, we still know little about the nature and impact of this ideal-employee factor, despite its likely relevance for both the personality tests' construct and criterion related validity (Morgeson et al., 2007a).

The current study delves into this last finding in the response distortion literature. Specifically, the goal of the study is to address three fundamental questions about the nature and workings of the ideal-employee factor (Schmit & Ryan, 1993): First, we present evidence for the ideal-employee factor as an indicator of applicant response distortion. Second, we draw on the literature on cognitive schemata (Fiske & Taylor, 1991) to explain how the ideal-employee factor may come about. Finally, we use this literature combined with the literature on candidates' ability to identify criteria (Kleinmann, 1993) to examine the ideal-employee factor's influence on the criterion-related validity of personality tests, thus paving the way for a theoretically grounded understanding of both the basic nature and the consequences of the ideal-employee factor for the criterion-related validity of personality tests during personnel selection (see Figure 1).

The Ideal-Employee Factor

Empirical Research

Schmit and Ryan (1993) first identified the sixth factor in applicant samples of personality data. They used invariance tests for comparing the factor structure of Big Five data gathered for research and for selection purposes. In line with the voluminous research on the Big Five, five factors seemed to underlie the personality data collected for research purposes. Conversely, in the applicant data gathered in a selection context, there was also evidence for a sixth factor. In particular, all items loaded on their respective Big Five factor and a sixth factor emerged that included loadings across different desirable personality dimensions, indicating applicants to be conscientious and productive hard workers, highly likable, courteous, thoughtful, considerate, organized, active, self-reliant, etc. Schmit and Ryan labeled this factor the *ideal-employee factor*.

It is important to stress that this sixth factor emerging on personality inventories during personnel selection does not detract from the large body of research on the construct-related validity of Big Five personality traits (Digman, 1990; Goldberg, 1990). In fact, all items still load on their designated Big Five factor. However, in addition to these expected loadings, some items also cross-load on a sixth factor. This factor is composed of items that assess *different* personality traits. The fact that self-assessments of such traits differ in the degree to which they load on the sixth ideal-employee factor suggests that this factor is not merely common method variance.

Since Schmit and Ryan's (1993) original study, indications for this sixth factor or high scale-intercorrelations have emerged in several applicant datasets (Cellar, Miller, Doverspike, & Klawsky, 1996; Collins & Gleaves, 1998; Ellingson et al., 1999; Pauls & Crost, 2005; Topping & O'Gorman, 1997; and conference papers by Biderman & Nguyen, 2009; Biderman, Nguyen, Mullins, & Luna, 2008; Burns & Christiansen, 2007; M. A. Smith et al., 2001; Van Iddekinge, Raymark, Eidson, & Putka, 2001; see also Wiggins, 1959). While different scholars have used

different labels such as "faking factor" (Pauls & Crost, 2005), "faking ability" (Biderman &

Nguyen, 2004; Clark & Biderman, 2006; Wrensen & Biderman, 2005), or "ideal-applicant self-

presentation" (Morgeson et al., 2007a) for the same phenomenon, we follow the original work by

Schmit and Ryan by using the term "ideal-employee factor". Our first hypothesis essentially

replicates prior field studies in a simulated selection process:

> *Hypothesis 1*: A six-factor solution, modeling a common ideal-employee factor in addition
>
> to the Big Five personality dimensions, will provide a better fit to personality data
>
> gathered in a simulated selection context than a five-factor solution.

Underlying Theory

The theoretical rationale for the ideal-employee factor stems from a person-situation

interaction perspective (Fiske & Taylor, 1991; Mischel & Shoda, 1995, 1998). This perspective

argues that situational cues (e.g., items of a personality inventory and the context in which they

are completed) activate a series of mental representations or schemata within respondents –

cognitive structures that integrate memories, affective reactions, and inferred traits for behaviors

in a specific domain (Markus, 1977). These highly integrated knowledge structures, including the

information on how to apply this knowledge (Holden, Kroner, Fekken, & Popham, 1992),

provide the cognitive context in which incoming social information is processed and courses of

action are undertaken (Mischel & Shoda, 1995, 1998).

When completing personality tests, the specific schemata activated among respondents

might differ depending on the setting: Low-stakes settings such as the voluntary participation in

anonymous research likely trigger a "stranger-description" frame of reference (Schmit & Ryan,

1993), a self-referenced evaluation of the fit between the item and the dominant self-schema

(Holden et al., 1992) that depicts to a stranger what that stranger could know about oneself.

Providing self-reports of their dominant behavioral tendencies and thus their personality is

reflected in the clear Big Five factor structures emerging from personality ratings under such conditions (Digman, 1990; Goldberg, 1990).

 More evaluative settings such as personnel selection, however, may activate a rather different schema, namely an ideal-employee schema (Schmit & Ryan, 1993) which integrates knowledge about the presumed traits of a qualified candidate for the focal position (Holden et al., 1992). After all, in a selection context applicants' primary concern is less one of accurate self-report but of emerging as the ideal employee for the given position (J. Hogan et al., 2007). Whether caused by a conscious decision to distort responses in order to meet the requirements of an 'ideal employee', or caused by an unconscious process by which applicants envision themselves successfully fulfilling the requirements of their desired position, the result is that applicants tailor their answers to their assumptions about the ideal applicant (D. B. Smith, Hanges, & Dickson, 2001). They conduct a selective memory search for and/or present themselves more favorably on those traits deemed relevant (Kunda & Sanitioso, 1989; Sanitioso, Kunda, & Fong, 1990), resulting in personality ratings that may reflect not only their overall personality but also their perceptions of what constitutes adequate attributes of the ideal applicant (Guion, 1965; Dipboye in Morgeson et al., 2007b). All of this seems to result in an additional factor emerging in factor analyses of personality ratings of applicant data.

 When thinking of what constitutes an ideal applicant, most of the big-five dimensions will show at least some social desirability, i.e., one can assume employers in general to appreciate new hires who are, among other things, friendly, happy to communicate freely and to learn new things. In comparison, however, some personality dimensions may be particularly desirable. For instance, employers will likely search particularly for employees with high work-ethics, i.e., conscientiousness (Huffcutt, Conway, Roth, & Stone, 2001), given that conscientiousness is a particularly good predictor of performance (Barrick & Mount, 1991). In addition, organizations

may strive to find candidates who promise a certain level of emotional stability or steadiness in the face of whatever adversities they may encounter during their work (Huffcutt et al., 2001).

Consequently, perceptive applicants might present themselves favorably particularly on these two dimensions. In line with this notion, past research on response distortion among true applicants has shown that score inflation is usually highest on Conscientiousness and Emotional Stability (Birkeland et al., 2006) with smaller inflation emerging for the remaining three dimensions. Combining these considerations, we offer:

Hypothesis 2: Measures of Emotional Stability and Conscientiousness will load higher on the ideal-employee factor emerging under simulated applicant conditions than do measures of the remaining three Big Five personality dimensions.

Applicants' Assumptions of the Criteria Targeted in Selection

Whether or not they always are aware of it, people strive to control how others perceive them during social interactions (Baumeister, 1982). This striving should be particularly strong when a positive impression may lead to a desired job offer (R. Hogan, Hogan, & Roberts, 1996; R. Hogan & Shelton, 1998; Motowidlo, 1999). Yet, such contexts do not necessarily reveal what "positive" entails and it is therefore up to applicants to identify the criteria within the particular selection situation (Kleinmann, 1993). Thus, some individuals' schemata (Holden et al., 1992) about the requirements of the job in question may be more accurate than those of others.

If the schema activated in a selection context reflects applicants' knowledge about the presumed traits of a qualified candidate for the focal position, this implies that candidates can possibly identify and know the targeted criteria in a given context. Kroger and Turnbull (1975) found that participants of a laboratory experiment could faithfully reproduce the personality profile associated with a culturally unambiguous social role, but that their 'role faking' was less successful for a more ambiguous social role. As soon as participants were provided with more

information, however, allowing them to gain a more accurate perspective of this formerly ambiguous role, their simulated profiles grew considerably more accurate.

Beside situational transparency, however, also candidates differ in their ability to identify the relevant schemata or, in other words, in their ability to identify the criteria (ATIC) relevant for the current position. Specifically, recent research has shown that identifying the criteria targeted in a selection situation is not an easy feat to accomplish. On average, candidates usually identify only about a third to half of the requirements inherent in different selection situations such as assessment centers (Kleinmann, 1993; Preckel & Schüpbach, 2005), structured interviews (Melchers et al., 2009), or integrity tests (König, Melchers, Kleinmann, Richter, & Klehe, 2006). Rarely is a required performance dimension accurately identified by all participants of a given study (Kleinmann, 1993). Additionally, individuals differ both substantially and reliably in their ATIC, i.e., the degree to which they can discern what is required of them (e.g., Kleinmann, 1997a; König, Melchers, Kleinmann, Richter, & Klehe, 2007). Finally, candidates' ATIC consistently predicts their performance in (Kleinmann, 1993; Melchers et al., 2009) as well as across different performance situations (König et al., 2007). In essence, being able to interpret the cues that indicate what is required and used for evaluating performance in a given situation helps candidates to adjust their responses and to thereby be more likely to succeed in that context.

Consequently, we assume that candidates' ATIC also enables them to conceptualize the ideal employee profile for a given selection procedure. Candidates then use this conceptualization as a template to complete the items of a personality inventory. They inspect each item (see Knowles, 1988) and compare it to the ideal employee profile that they constructed via their ATIC. If an item confirms their construal of the ideal employee, they provide a more concordant self-presentation. This item-by-item comparison process could explain why some items load more heavily onto the sixth factor than others.

Hypothesis 3a: Participants' ATIC will be related to the level to which their self-reported

personality assessments in a simulated selection process reflect an ideal-employee factor.

Predicting Job Related Performance

Besides possibly influencing candidates' personality test responses, ATIC is a reliable

predictor of performance in and across different assessment situations (Kleinmann, 1993, 1997a;

König et al., 2007; Melchers et al., 2009). The relationship between ATIC and job-related

performance has been explained by the notion that ATIC is a measure of social perceptiveness,

i.e., the cognitive understanding or savvy with which people accurately read particular situational

demands (Ferris, Perrewé, & Douglas, 2002). In line with this assumption, ATIC correlates

positively with self-reported social skills (Schollaert & Lievens, 2008) and with performance in a

video-based social judgment test (Kleinmann, 1997b), besides showing modest correlations with

general mental ability or sub facets thereof (e.g., König et al., 2007; Melchers et al., 2009).

Social perceptiveness has repeatedly been proposed to account for a possible positive link

between response distortion and performance (R. Hogan & Shelton, 1998; J. Hogan et al., 2007;

Hollenbeck in Morgeson et al., 2007b). Feeding into the debate about whether response distortion

– indicated via the emergence of an ideal employee profile – is negatively related, unrelated, or

even positively related to applicants' work related performance (Komar, Brown, Komar, &

Robie, 2008), our basic premise is that both candidates' processing of the personality test's items

as well as their performance on job related criteria is a function of their ATIC. ATIC might thus

be positively related to the emergence of the ideal-employee factor, assuming that particularly

applicants with a high ATIC distort their responses towards an ideal employee profile. In line

with prior research, ATIC should further predict measures of job-related performance. It follows

that ATIC serves as a common source of both the ideal-employee factor and job-related

performance. Another corollary is that the relationship between the ideal employee factor and

job-related performance would be positive only until candidates' ATIC is controlled.

Hypothesis 3b: Participants' ATIC will predict their job-related performance.

Hypothesis 3c: Participants' ATIC will serve as a common source variable between candidates' scores on the ideal-employee factor and their job-related performance: Without the consideration of ATIC, the ideal-employee factor will be positively related to candidates' job-related performance, whereas this relationship will become negligible when controlling for ATIC.

Rival Explanations

The last hypothesis implies that response distortion may not be such a negative thing after all and the ideal-employee factor might, in fact, predict job-related performance (R. Hogan & Shelton, 1998). This mirrors ideas that the ability to present oneself favorably on personality tests may be related to positive self-presentation skills (e.g., R. Hogan, 1991; Marcus, 2003, 2009) or to a functional awareness of social norms in line with Allport's (1937, p. 465) view that "every response is determined in part by adaptive performance to the specific demands of a situation". Similarly, Viswesvaran and Ones (1999, p. 207) suggest that *fakability*, the ability to fake, is a potentially useful individual difference variable: "for example, to the extent that fakability reflects social intelligence or some form of adaptability, individual differences in fakability may contribute to explaining successful job performance".

ATIC may thus not be the only possible variable accounting for common variance in the ideal-employee factor and performance. To discount possible rival explanations, we controlled for several variables. First, past research has repeatedly proposed that individuals high on general mental ability (GMA) have better test-taking skills (or "test smarts") and hence can more successfully distort personality items (Stricker, 1969; but see Mersman & Shultz, 1998; Ones et al., 1996), a proposition partially supported by Christiansen et al. (2005) and Vasilopoulos,

Cucina, Dyomina, Morewitz, and Reilly (2006). Also, prior research found moderate relationships between ATIC and GMA (Melchers et al., 2009; Preckel & Schüpbach, 2005).

Second, ATIC has potential conceptual overlap with self-monitoring; that is, the extent to which individuals monitor, adjust, and control their behavior based on how it is perceived by others (Snyder, 1974). The conceptual difference between ATIC and self-monitoring is their basic nature as either a primarily ability-related or motivational variable. Gangestad and Snyder (2000, p. 547) concluded that self-monitoring "relates to status-oriented impression management motives" and usually self-monitoring is measured via relatively generalized self-ratings. ATIC, in contrast, is measured via a context specific perceptual ability test (see Methods).

Study Summary

In short (Figure 1), we expect that in personnel selection situations, Big Five personality items will load not only on their designated personality construct but also on a common latent ideal-employee factor (Hypothesis 1), with Conscientiousness and Emotional Stability showing particularly high loadings (Hypothesis 2). Candidates' ATIC will predict both this ideal-employee factor (Hypothesis 3a) and candidates' performance (Hypothesis 3b), accounting for an otherwise significant link between the ideal-employee factor and performance (Hypothesis 3c).

This study aims to illuminate the nature and cause of the ideal-employee factor as an indicator of response distortion. It will also provide a needed test of whether response distortion is positively, negatively, or not related to measures of performance (Komar et al., 2008). In doing so, this study might provide a conceptual explanation for why the criterion related validity of personality assessments seems to be no worse in applicant situations than under more neutral conditions, despite the likely occurrence of response distortion (Barrick & Mount, 1996).

Methods

Setting

Given that the ideal-employee factor has up to now only been found in applicant samples, we needed to assess candidates' responses on the personality inventory in an ecologically valid application setting (e.g., Pauls & Crost, 2005). Also, the subsequent task performance needed to be relevant and engaging for participants (Mesmer-Magnus & Viswesvaran, 2006). At the same time, we needed a setting that allowed for the objective and reliable assessment of ATIC, independently scored – rather than a merely self-reported – against a priori criteria. For this purpose, performance needed to be assessed on pre-established dimensions that participants either would or would not identify correctly, using equivalent tasks across participants.

Consequently, the present research used an experimental protocol employed in earlier studies (Kleinmann, 1993, 1997a; Kleinmann, Kuptsch, & Köller, 1996; Klehe, König, Kleinmann, Richter, & Melchers, 2008; Melchers et al., 2009) in the form of a simulated selection process organized by the psychology departments and career centers of two universities and a local branch of the German Federal Employment Office. Target participants were university graduates who were applying for a job or would soon do so. This setting offered standardized conditions for assessing the relevant variables, namely paper-and-pencil self-reports of the Big Five and self-monitoring, together with GMA and performance on various work simulations, as well as ATIC pertaining to these simulations.

Development of the procedures. The process focused on the position of a management trainee, as within the German context such a position represents a realistic and attractive job for university graduates from diverse academic backgrounds. Based on a job analysis for management trainees, subject matter experts rated the following three dimensions to be most conceptually independent from one another and most assessable during high- and low-fidelity work simulations (see Kleinmann, 1997a; Kleinmann et al., 1996, for a full description of this procedure): Planning was defined as prioritizing tasks, making plans for tasks and projects,

making appointments in due time and allocating tasks. Leadership was defined as striving for and

taking on responsibility for tasks and groups, coordination of teams and arguing one's point of

view within a group. Finally, Cooperation was defined as consideration of others' needs and

assisting with others' problems, as well as being prepared to compromise and to mediate between

diverging interests and points of view.

Job-related performance. We assessed job-related performance using four high-fidelity

and 24 low-fidelity work simulations that allowed the standardized observation of participants'

performance over two days (e.g., Klehe et al., 2008; Kleinmann, 1997b). The high-fidelity

simulations were chosen from a list of eight simulations deemed usable for assessing the chosen

dimensions Planning, Leadership, and Cooperation by three personnel experts. For making this

choice, twelve trained observers learned about the three targeted dimensions, performed each

simulation themselves and rated the three dimensions' relevance for good performance in each

simulation. These ratings lead to the choice of an organizing task in which each participant had to

plan and schedule numerous events, a business presentation, and two leaderless group discussions

simulating different organizations' board meetings. One of these discussions, regarding the

development of a new strategy in the face of changed market-conditions, used assigned roles with

opposing interests (finance, personnel, production, and sales), while the other discussion (on

whether to extend domestic production and/or move it abroad) had no assigned roles.

The low-fidelity work simulations were chosen from a set of 34 past- and future-oriented

low-fidelity simulations (Motowidlo, Dunnette, & Carter, 1990) collected by two industrial and

organizational (I/O) Psychology doctoral students from validated scenarios used in field settings

or earlier studies. The reason for including not only high- but also low-fidelity simulations is that

low-fidelity simulations can cover a relatively broad range of relevant work-related situations in

relatively little time. For each simulation, behavioral scoring guides provided anchors for poor

(1), acceptable (3), and outstanding (5) answers. Ten I/O Psychology Masters students evaluated the understandability of items and anchors, and rated how well each scenario measured the targeted as well as various other dimensions. Only scenarios that clearly addressed one of the intended dimensions but no second and/or non-intended dimension were chosen for the final set of situations, resulting in twelve past- and twelve future-oriented scenarios, with four scenarios per format covering each of the three targeted dimensions. A sample low-fidelity scenarios for assessing *Cooperation* is: "Imagine you'd been assigned a new and fascinating project which requires considerable effort and attention. As one of your old projects has not yet been completed, management has given responsibility of this old project to someone else who had previously been uninvolved in it. How would you handle this situation?"

  Raters. The raters, most of whom were I/O Psychology Master's students, participated in a one-day observer training session during which they learned about the different simulations and targeted dimensions. Raters learned about typical rating errors and discussed each simulation and the behavioral anchors in order to achieve a consistent frame-of-reference for rating participants' performance (Latham & Wexley, 1994; Woehr & Huffcutt, 1994). Raters did not receive information concerning the objectives of the study.

Participants

  Participants were recruited via flyers and the universities' career centers to take part in a professional hands-on applicant training program. They knew that the training would consist of a simulated selection process and they participated in the study in order to learn about different types of tasks that are usually employed during personnel selection and to receive individual feedback on their own behavior and performance. We explicitly addressed recent or soon-to-be university graduates in order to ensure high participant interest and to enhance the generalizability of our results. Of the 149 participants (48% men), 39% held a Master's degree,

36% an undergraduate degree and 25% were still completing their undergraduate studies. On average, participants were 27 years old ($SD = 4.6$) and had studied for 4 years ($SD = 2.1$) in business administration (43%), other social sciences (27%), and natural sciences (30%). Most participants indicated that the application situation appeared realistic (93%) and that they put themselves into the position of an applicant (91%) during the simulated process.

Procedure

In order to enhance the realism of the simulations and to give participants some indication regarding the requirements of the tasks, participants prepared a written application prior to attending the simulated selection process. For this purpose, they had received a job advertisement for the management trainee position (see Appendix A). This advertisement included subtle information about the three relevant dimensions, such as the requirement for job incumbents to take responsibility (i.e., *Leadership*), without ever being informed directly about the dimensions targeted in the work simulations. Participants also received some more information at the start of the actual simulation in the form of a written report of a current trainee (Appendix B). The subsequent two day simulation centered around this position just as if this was the job that participants were actually applying for. Participants received an individualized schedule informing them where and when to attend which test or simulation. They were also provided with standard instructions before each of the assessment tasks, as well as being given feedback and being debriefed after all the study tasks had been completed.

Job-related performance. Four raters observed and evaluated participants during the two group discussions and two raters evaluated participants during the organizing task, the presentation, and the 24 low-fidelity simulations. Job-related performance on each was always scored from 1 (*poor*) to 5 (*outstanding*). The average interrater agreement (i.e., the average correlation between the raters) for the overall performance ratings was .80 and .78 for the two

group discussions, .85 for the business presentation and .95 for the organizing task, which is typical for these types of tasks (Collins et al., 2003). The average interrater agreement for performance in the low-fidelity simulations (averaged across all 24 items) was .92, which again is comparable to meta-analytically derived values (cf. Conway, Jako, & Goodman, 1995).

ATIC. We assessed participants' ATIC using the procedure outlined in König et al. (2007) and Melchers et al. (2009), a procedure aimed at measuring participants' ATIC as a true measure of ability. Specifically, after each high-fidelity simulation and the two sets of low-fidelity simulations, participants learned that the following task was done only for research-purposes. They then received an open ended questionnaire that asked them "In the previous situation, you possibly thought about what the observers were assessing. What assumptions did you have during the situation about what the simulation was intended to assess?" Like in the studies of König et al. and Melchers et al., participants could write down up to two hypotheses for every low- and up to six hypotheses for every high-fidelity simulation.

After the completion of the two-day assessment but before receiving feedback about their performance, participants learned about different performance dimensions that frequently play a role in both job and selection situations. Besides the three dimensions actually observed, these dimensions included the three bogus dimensions job knowledge and experience, self-confidence, and acquisition and handling of information (Huffcutt et al., 2001). Participants also received a list of behavioral examples for each dimension. After having read this list they received back the questionnaires in which they had written down their own assumptions subsequent to each simulation. They now indicated for each of their assumptions whether it corresponded to any of the dimensions listed, as well as the strength of this correspondence on a scale from 1 (= *fits somewhat*) to 4 (= *fits completely*). Participants could also indicate that an assumption did not correspond to any of the dimensions.

Finally, the experimenters computed how correctly participants had identified each of the

dimensions actually observed. For this purpose, they checked whether the participant had

identified every dimension that had actually been observed. If participants had not indicated the

dimension to be representative of one of their own assumptions, then they received an ATIC

score of 0 for this dimension. If participants had noted that the dimension represented one of their

own assumptions, then participants' rating of the strength of the fit between their assumption and

the dimension was assigned as the ATIC score for this dimension. In the case of ties (several

assumptions being linked to the same dimension), we used the highest strength of fit rating as the

score. This scoring procedure resulted in ATIC values ranging from 0 (*correct dimension not*

identified) to 4 (*assumption fully fit the correct dimension*) for every dimension observed in every

exercise.[1] As expected, a scree-plot suggested the existence of one factor explaining the variance

in ATIC scores, and a measurement model with a single factor resulted in an acceptable fit,

$\chi^2(594) = 784.49$, $\chi^2_{/df} = 1.32$, RMSEA = .05. The ATIC measure's internal consistency was .84.

Big Five personality factors. Participants responded to the 60-item German NEO-FFI

(Borkenau & Ostendorf, 1993). According to the test authors, internal consistencies usually range

from .72 (Agreeableness) to .86 (Conscientiousness).

Control variables. In order to assess participants' GMA, participants completed the

following six subtests: Figures, Matrices, Analogies, Cubes, Sentence Completion, and

Similarities from the IST 2000 (Amthauer, Brocke, Liepmann, & Beauducel, 1999), a widely

used and valid German cognitive ability test (Hülsheger, Maier, Stumpp, & Muck, 2006).

Amthauer et al. (1999) reported an internal consistency of .88.

We assessed self-monitoring at the end of the simulated selection procedure with the ten-

item social comparison scale, a widely used German self-monitoring scale by Nowack and

Kammer (1987). Items had been developed to identify the degree to which individuals pay

attention to cues about adequate social behavior. A sample item is "*I'm very interested in other people's opinion about my behavior.*" In previous investigations, coefficient alpha for this scale had been around .74 (Mielke & Kilian, 1990; Nowack & Kammer, 1987).

Results

Preliminary Analyses

Table 1 presents the means, standard deviations, internal consistencies, and correlations between variables. As expected, all measures showed acceptable internal consistencies and candidates' responses to the Big Five personality dimensions showed meaningful correlations with their performance in the high- and low-fidelity work simulations, a finding consistent with prior studies (Barrick & Mount, 1991).

Additional analyses served to test whether our simulated selection situation showed ecologic validity in eliciting similar findings as reported in earlier studies on personnel selection: This research had found that applicants reported significantly higher personality scores than would normally have been expected, with score inflation being highest for Conscientiousness and Emotional Stability (Birkeland et al., 2006). In order to test whether the same was true among the current participants, we compared participants' scores to the scores of the norm-population. Reported in the official NEO-FFI handbook, the norm-population combines the responses of 11.724 representative men and women from 50 non-clinical personality studies conducted under anonymous/low-stakes settings in Germany, Switzerland, and Austria (Borkenau & Ostendorf, 1993). It can thus serve well as an estimate of the average scores that one would have expected under non-selection conditions.

This comparison indeed revealed inflated personality scores among the participants of the current study: Participants gave themselves significantly higher ratings than norm-values would suggest on fifty-two of the sixty items. This difference was most pronounced for the two

dimensions Conscientiousness (d = 1.10) and Emotional Stability (d = .92) with participants on average reporting scores that would place them into the 87th and 82nd percentile of the norm population, respectively (Cohen, 1988, p. 22). In summary, these data mirror earlier findings on applicant response inflation during personnel selection (Birkeland et al., 2006) and thus further suggest that participants took the simulated selection process in this study seriously and acted like they would have done during an actual selection situation.

Hypothesis 1

Hypothesis 1 proposed that the Big Five personality dimensions would load not only on their respective latent personality dimensions, but additionally on one common ideal-employee factor. Schmit and Ryan's (1993) discovery of the ideal-employee factor under applicant conditions was based on exploratory factor analyses results. Given earlier replications of this factor under applicant conditions, however, and consistent with subsequent research (e.g., Cellar et al., 1996; Pauls & Crost, 2005), we developed *a priori* hypotheses and adopted a confirmatory factor analytic approach which relied on the comparison of two competing models.

Model 1 assumed the presence of the five correlated latent variables Emotional Stability, Conscientiousness, Agreeableness, Extraversion, and Openness to Experience. This model represents the traditional measurement model underlying personality data in non-evaluative situations. In order to prevent analyses with an unwieldy number of items (60 items loading on five factors), each of the five factors was defined not by twelve individual items but by three parcels of items with each parcel being the average value of four of the scale's items, with items being assigned to parcels on a random basis (cf. Hall, Snell, & Foust, 1999). Every one of the 15 resulting parcels was specified to load only on its own latent Big Five factor. The five dimensions were allowed to covary with one another.

Model 2 mirrored Model 1 except for an additional latent variable representing the ideal-

employee factor. This model represents the measurement model underlying personality scores

obtained during personnel selection situations. As done in earlier studies on the ideal-employee

factor (e.g., Cellar et al., 1996; Pauls & Crost, 2005), the additional factor was allowed to impact

every one of the 15 item-parcels, yet was uncorrelated with the latent personality factors in

Model 2 (see Cellar et al., 1996, p. 699).[2]

To test how well the nested Models 1 and 2 fit the data, we used the overall model χ^2 and

the $\chi^2/_{df}$ ratio, which should be below 3 and generally as low as possible (Byrne, 1994, 1998). An

acceptable fit is further indicated by an incremental fit index (IFI), a Tucker-Lewis index (TLI)

and a comparative fit index (CFI) of at least .90 and preferably higher, as well as by a root mean

square error of approximation (RMSEA) of at most .08 and preferably lower (Byrne, 1998). In

order to support Hypothesis 1, Model 2 further needed to show a better fit than Model 1, as well

as significant paths between the proposed ideal-employee factor and the measurement parcels.

Both models fit the data reasonably well (Table 2). In comparison, however, Model 2 fit

the data better than Model 1, as is evident from the significant reduction in χ^2, $\Delta\chi^2(15) = 64.32$, p

< .01, and the substantial increase in CFI (Cheung & Rensvold, 2002). Additionally, the ideal-

employee factor significantly influenced most of the individual parcels. In sum, this replicates

earlier findings on the emergence of the ideal-employee factor under applicant conditions.[3]

Hypothesis 2

Hypothesis 2 proposed that Conscientiousness and Emotional Stability measures would

load significantly higher on the ideal-employee factor than would measures of the remaining

personality dimensions. In Model 2, Conscientiousness and Emotional Stability parcels loaded in

average .67 and .53, respectively, on the ideal-employee factor. The average loadings of the

remaining personality parcels appeared somewhat lower (.34 for Agreeableness, .21 for

Extraversion, and .17 for Openness parcels, respectively). To test whether this difference was

statistically significant, we ran a second set of model-comparisons (Models 3a to 3n). In these comparisons, we used the highest factor loading on the ideal employee factor identified in Model 2, a loading belonging to one of the Emotional Stability parcels, as a benchmark against which to test the remaining factor loadings on the ideal employee factor. Particularly, we constrained the factor loadings of different item-parcels to be equal to this factor loading. We expected this procedure to lead to no decrement in model fit for the remaining Emotional Stability and Conscientiousness parcels, thus showing that their factor loadings are comparable to the highest factor loading identified. For the factor loadings belonging to Extraversion, Agreeableness, and Openness parcels, however, we did expect a significant decrement in fit, thus showing that these factor loadings were significantly smaller than the highest factor loading identified in Model 2.

Results (Table 2) indicated that all Conscientiousness parcels, two of the three Emotional Stability parcels, and one Extraversion parcel loaded equally highly on the ideal-employee factor as did the highest loading parcel, while the last Emotional Stability parcel, just as the remaining parcels belonging to Extraversion, Agreeableness, and Openness to Experience, showed significantly lower loadings. In sum, this finding largely supports the notion that the ideal-employee factor particularly impacts responses to items that ask candidates to describe themselves as conscientious, goal driven, systematic and hard working, as well as self-reliant, composed, cheerful and stress resistant. Additionally, a parcel of extraversion items asking candidates to describe themselves as energetic, optimistic, and active showed a comparable loading onto the ideal-employee factor. This largely supports Hypothesis 2.

Additional Analyses on the Ideal-Employee Factor

In order to further validate the ideal-employee factor as an indicator of candidate response distortion, we compared findings associated with the ideal employee factor with those associated with personality test score inflation on both an item and a person level. After all, score inflation is

the most persistent and most troubling finding about response distortion (J. Hogan et al., 2007) with some items being particularly vulnerable to inflation (Birkeland et al., 2006) and some candidates inflating scores more than others (Schmitt & Oswald, 2006). If score inflation and the ideal employee factor address the same phenomenon, there should be a high positive correlation between the two. We tested this assumption on both the person- and the item-level.

On the person-level, this implies that the higher a participant scores on the ideal-employee factor, the more this person should also inflate his or her personality scores when compared to scores expected under non-applicant conditions. We again used the norm-values provided in the NEO-FFI test manual (Borkenau & Ostendorf, 1993) on the participant level as "honest" scores. Comparing these values to the values obtained in the current study, we estimated (Δ) each participant's likely score inflation across items. This estimate is admittedly rough, given that it does not differentiate between dimensions and that we don't know each participant's 'honest' score – i.e., while there is no reason to assume that our sample, in average, differs from the norm-population on their honest NEO-FFI scores, we cannot conclusively ensure that any one participant ascribing themselves high scores does so because he or she engages in score inflation or because this person actually has an above-average scoring personality profile.

Then, we imputed each person's ideal-employee factor-score via structural equation modeling. When we correlated these ideal-employee factor-scores with participants' estimated score inflation, the resulting r of .70 ($p < .01$) suggested that particularly participants scoring high on the ideal-employee also showed inflated scores when compared to the test's norm-population.

On the item-level, the assumption that score inflation and ideal employee factor represent the same phenomenon implies that the higher an item loads on the ideal-employee factor, the more this item should also is be inflated when compared to the item-score expected under non-applicant conditions. We now used the item level NEO-FFI norm-values (Borkenau & Ostendorf,

1993) as "honest" scores and compared them with the values obtained in the current study via two-sample t-tests, thus calculating each item's score inflation (effect size d) across participants. A particular advantage of this approach, compared to the person-centered approach above, is that each item's 'honest' value should be well represented by the NEO-FFI norm-value and we see no conceptual reason – except for score inflation – to believe that any one item's average score across participants should meaningfully differ from the score represented for that item in the norm values. Next, we estimated each item's loading onto the ideal-employee factor via structural equation modeling.[4] When we correlated the 60 items' loadings onto the ideal-employee factor and the d-value of the inflation for each item, we found an r of. 89 ($p < .01$). This indicates that an item's loading on the ideal-employee factor stands in a nearly perfect linear relationship with the degree to which this item is being inflated. In sum, these results support the notion that the emergence of the ideal employee factor is another indicator of applicants' response distortion.

Hypothesis 3

The final set of hypotheses addressed the role of participants' ATIC in linking the ideal-employee factor and participants' performance. In particular, Hypothesis 3 proposes that the ideal-employee factor would be a function of participants' ATIC (Hypothesis 3a), and that ATIC would also predict participants' performance in the different high- and low-fidelity work simulations (Hypothesis 3b). ATIC would thus account for an otherwise positive relationship between the ideal-employee factor and performance (Hypothesis 3c).

Again, we tested this assumption via a model comparison procedure. The proposed Model 4a (Figure 1) assumed ATIC to predict both the ideal-employee factor (Hypothesis 3a) and performance (Hypothesis 3b), while we also controlled for the impact of the Big Five personality dimensions, self-monitoring and general mental ability on performance and the impact of self-monitoring and general mental ability on the emergence of the ideal-employee factor. In this

model, the direct path from the ideal-employee factor to performance should be nonsignificant (Hypothesis 3c). This path should, however, become significant (Model 4b) as soon as the two paths from ATIC to the ideal-employee factor and to performance were deleted from the model.

In order to support Hypothesis 3, four conditions needed to be met: First, the two paths from ATIC to the ideal-employee factor (Hypothesis 3a) and performance (Hypothesis 3b) should be significant in Model 4a. Second, the deletion of these paths in Model 4b should significantly decrease the model's fit. Third, the ideal-employee factor should show no significant link to performance in Model 4a, yet should well show such a link in Model 4b.

As expected, ATIC showed positive relationships with both the ideal-employee factor, $\gamma = .54$, $p < .01$, and performance, $\gamma = .39$, $p < .05$ in Model 4a, while the direct path from the ideal-employee factor to performance was non-significant, $\gamma = .03$, $p = .86$. As soon as the two paths from ATIC to the ideal-employee factor and performance were deleted in Model 4b, however, the ideal-employee factor indeed predicted performance, $\gamma = .29$, $p < .01$. When comparing Models 4a and 4b (Table 3), Model 4a created a significantly better fit to the data, as is evident from the lower χ^2, $\Delta\chi^2(2) = 27.83$, $p < .01$, and higher CFI, indicating a substantial increase in model fit (Cheung & Rensvold, 2002). Thus, Hypothesis 3 was supported, while neither self-monitoring nor general mental ability contributed to the prediction of the ideal-employee factor. In summary, Model 4a (see Figure 2 for the full final model) accounted for 55% of the variance in performance and for 28% of the variance in the ideal-employee factor.

Discussion

Applicant response distortion has been an enduring concern in the literature on personality assessment (Allport, 1937; Griffith & Peterson, 2006; Rosse et al., 1998; Viswesvaran & Ones, 1999; Whyte, 1956). The literature is divided as to whether successful response distortion should be considered a useful skill (Viswesvaran & Ones, 1999; Murphy in

Morgeson et al., 2007b), inconsequential (Ones et al., 2007), or appalling (Campion in Morgeson et al., 2007b; Tett & Christiansen, 2007), though few studies have directly tested the distortion-performance relationship (Komar et al., 2008).

The current study informs this debate by tackling some individual differences underlying response distortion during personality assessments. We know that the criterion-related validity of personality scales does not suffer in applicant samples despite the occurrence of response distortion. Building on the literature on cognitive schemata (Fiske & Taylor, 1991), we sought to demonstrate that the ideal-employee factor found to underlie ratings of theoretically unrelated personality constructs in applicant samples is an indicator of applicant response distortion, and to explain how and why the ideal-employee factor may be linked to measures of performance.

Similarly to earlier studies (e.g., Collins & Gleaves, 1998; Ellingson et al., 1999; Pauls & Crost, 2005; Schmit & Ryan, 1993; Topping & O'Gorman, 1997; Van Iddekinge et al., 2001; Zickar & Robie, 1999), our data suggest that personality assessments under applicant conditions bring about the emergence of an ideal-employee factor spanning across different and conceptually unrelated personality dimensions. The notion that this factor is an indicator of applicant response distortion is supported by internal indications of construct validity. Our data supported earlier findings (Pauls & Crost, 2005; Van Iddekinge et al., 2001) that the ideal-employee factor particularly impacted measures of Conscientiousness and Emotional Stability, even though other dimensions were affected as well. Apparently, being hard-working and stress resilient fits the general stereotype of the ideal employee. Additionally, a comparison with each item's norm-value indicated a strong relationship between item score inflation and the item's loading on an ideal-employee factor. And finally, the more participants inflated their scores compared to representative norm-values, the higher they also seemed to score on the ideal-employee factor. In sum, it seems that the ideal-employee factor emerging in personality

assessments under applicant conditions is an indicator of applicant response distortion (see also Biderman & Nguyen, 2009, for similar results).

The second contribution of this study lies in providing evidence for the importance of ATIC as a key individual difference variable. Scholars have long called for additional theorizing and research on the process and outcomes of applicant response distortion in order to learn why distortion does not diminish the criterion-related validity of personality assessments (Komar et al., 2008; Ones et al., 2007; Viswesvaran & Ones, 1999). ATIC not only predicted performance as it had in prior studies (Kleinmann, 1993; König et al., 2007; Preckel & Schüpbach, 2005), it also opened the black box of the nature of applicants' "ability to fake" proposed in most models of response distortion: While models often include assumptions about the cognitive factors underlying of applicants' ability to distort responses (Marcus, 2003; McFarland & Ryan, 2000; Snell, Sydell, & Lueke, 1999), such assumptions have rarely been tested directly (Christiansen et al., 2005). The present findings begin to establish ATIC as a key individual difference variable that accounts for substantial variance in the ideal-employee factor and its link to performance in different job-related situations. More generally, the ability to identify criteria (ATIC) is thus related to and may actually represent a substantial component of McFarland and Ryan's (2000, 2006) concept of "ability to fake". In contrast to the traditionally negative connotation associated with the term "faking", however, one may regard ATIC as a positive social skill (König et al., 2007) similar to Viswesvaran and One's (1999) concept of "fakability as a form of social intelligence." Indeed, given that response distortion seems inevitable and endemic (Morgeson et al., 2007b; Tett & Christiansen, 2007), an ability to do it well could constitute a form of insightful behavior (R. Hogan & Shelton, 1998; Morgeson et al., 2007b; Viswesvaran & Ones, 1999). Such an ability might support individuals in interpreting and reacting appropriately to both selection and job situations. Thus, the cognitive schemata apparently underlying "faked" personality scores

during personnel selection may contribute to the personality test's criterion related validity, an effect that can be explained by applicants' ATIC for successful performance.

Another contribution of this study is therefore the evidence of a positive relationship between the ideal-employee factor and participants' performance in diverse work-related simulations (see also Biderman et al., 2008). This is important in so far as a major concern for practitioners and scientists alike is whether faking on personality tests is positively, negatively, or not related to performance. Komar et al. (2008) demonstrated that the impact of faking behavior on validity is primarily determined by the faking-performance relationship. While the existence of an ideal-employee factor may complicate the construct validity of personality measures in a selection context, our results support earlier assumptions (R. Hogan & Shelton, 1998; Murphy in Morgeson et al., 2007a, 2007b; Ones et al., 2007; Viswesvaran & Ones, 1999) that it may not necessarily lower the criterion-related validity of such measures.

Limitations and Directions for Future Research

This study is not without limitations. We chose the setting for the current study in order to meet a number of different requirements; primarily, the assessment of all relevant variables, clearly defined performance dimensions, and standardized assessment conditions. At the same time, this setting also bears a number of potential disadvantages, most prominently the lack of an actual on-the-job-performance criterion. The performance proxy criteria we employed were various high- and low-fidelity work simulations. Such simulations are a good context to scrutinize personality-performance relationships because they simulate key job situations and require candidates to demonstrate behavior in these simulated work contexts.

In addition, our finding of a positive relationship between the ideal-employee factor and performance, combined with a poor ability of the original Big Five to predict performance, is in line with work by Biderman and colleagues (2008) who reported similar findings for supervisory

evaluations of work performance. Yet unlike such supervisory evaluations, the current criteria are likely to represent maximum rather than typical performance situations (e.g., Klehe, Anderson, & Viswesvaran, 2007; Sackett, 2007). Since the primary difference between typical and maximum performance situations lies in participants' higher motivation during maximum performance situations, motivational variables usually have a decreased and facets of ability an increased impact on performance under maximum performance conditions (Klehe & Anderson, 2007). Thus, research is warranted on whether in typical performance situations, the present results may underestimate the impact of personality on performance (see also ForsterLee, 2007; Marcus, Goffin, Johnston, & Rothstein, 2007) and potentially overestimate the predictive power of ATIC.

Second, performance simulations do not include a number of factors relevant in work settings (e.g., differences in job experience, leader-member exchange, day-to-day variations in performers' motivation) that could alter the effects we observed in this study. Subsequent research should thus replicate the current findings in field settings, even though doing so would probably lack some of the experimental controls enabled by the design of the present study. Another question may be whether ATIC maintains its predictive power over time or whether, particularly in routine jobs, ATIC may become less relevant as soon as the novelty of the job has worn off and performers know the ins and outs of the requirements posed towards them.

A third possible limitation is that we primarily focused on participants' ability to distort responses and less on their motivation to do so (Marcus, 2003; McFarland & Ryan, 2000). While the motivation to present oneself favorably is generally high across candidates in the "motivated" context of being a job applicant (Morgeson et al., 2007b), outright and conscious distortion likely makes up only "a small and insignificant part" of candidates' responses (Dipboye in Morgeson et al., 2007b, p. 692). In the current study, self-monitoring, the more motivational variable relating to status-oriented impression management motives (Gangestad & Snyder, 2000), showed no

meaningful impact on the ideal-employee factor, yet a full exploration of the construct validity of the ideal-employee factor that considers both ability-related and motivational faking-related variables should offer a fruitful avenue of future research. In the end, it will be interesting to learn how much of the response distortion observed during personnel selection in general and the ideal employee factor in specific is actually due to conscious distortion and how much of it is the plain and possibly even unconscious result of different schemata being used during personnel selection without respondents necessarily perceiving their answers as lacking in sincerity.

Further research on the construct validity of the ideal-employee factor might also be helpful in identifying boundary conditions to the positive relationship between the ideal-employee factor and performance. Komar et al. (2008) report that the distortion-performance relationship has a substantial impact on the personality test's validity, both directly and indirectly through its interaction with other parameters. As suspected in the traditional faking literature, distortion would diminish the test's criterion-related validity if distortion was unrelated or even negatively related to the performance criterion. That said, much of the prior research indicating the possibility of a negative relationship between response-distortion and performance has relied on intra-individual mean-comparison studies with "faking good" instructions that tend to be unrealistic (Blickle, Momm, Schneider, Gansen, & Kramer, 2009; Morgeson et al., 2007a).

At the same time, the current results were obtained with social criteria and an essentially socially-oriented explanation (ATIC as a measure of social perceptiveness). This does not exclude the possibility that such relationships might become non-significant or perhaps even negative when the criterion is less social in nature (e.g., software coding) and/or when the personality dimension distorted is of upmost importance for the job in question (e.g., emotional stability among soldiers, emergency personnel, or bomb disposal expert).

Relatedly, future research might further delve into the situational specificity of the ideal-

employee factor. The basic arguments underlying the proposition of an ideal-employee factor

stem, after all, from a person-situation interaction perspective. In the current study, we primarily

focused on the person (in the form of their ATIC), rather than the idiosyncrasies of the situation

involved, choosing a relatively general job-description that is applicable to candidates of many

educational backgrounds. Yet, some studies suggest that the factor structures for the ideal-

employee factor somewhat depend on the specificities of the position involved: Schmit and Ryan

(1993) found that particularly Conscientiousness, Agreeableness, and Extraversion items loaded

on this factor among a sample of job applicants seeking employment assistance from their

regional employment service, whereas Pauls and Crost (2005) found Emotional Stability,

Conscientiousness, and Extraversion to load on the sixth factor when participants were thinking

of applying for a management position. Agreeableness loaded additionally on that factor when

participants were thinking of applying for a nursing position. Thus, the content of the ideal-

employee factor might partially depend on the job of interest, with personality factors such as

Conscientiousness being both a consistent predictor of good performance and a consistent target

of applicant response distortion (Birkeland et al., 2006), while the response distortion observed

on other personality factors may depend on the specific requirements of the targeted job.

Conclusion

Building on the literature on cognitive schemata, the ideal-employee factor found to

underlie ratings of theoretically unrelated personality constructs in applicant samples appears to

be an indicator of applicant response distortion. Candidates' ATIC seems to drive the ideal-

employee factor and to account for the positive relationship observed between the ideal-employee

factor and performance: individuals with the ability to discern critical performance criteria are

also better at providing an ideal-employee profile on a personality inventory and at behaving in a

way consistent with this profile in a performance situation.

References

Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt.

Amthauer, R., Brocke, B., Liepmann, D., & Beauducel, A. (1999). *Intelligenz-Struktur-Test 2000 (IST 2000)*. Göttingen, Germany: Hogrefe.

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1-26.

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology, 81*, 261-272.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9*, 9-30.

Baumeister, R. (1982). A self-presentational view of social phenomena. *Psychological Bulletin, 91*, 3-26.

Biderman, M. D., & Nguyen, N. T. (2004). Structural equation models of faking ability in repeated measures designs, *19th annual conference of the Society for Industrial and Organizational Psychology*. Chicago, IL.

Biderman, M. D., & Nguyen, N. T. (2009). Measuring faking propensity, *24th Annual Conference of The Society for Industrial and Organizational Psychology*. New Orleans, LA.

Biderman, M. D., Nguyen, N. T., Mullins, B., & Luna, J. (2008). A method factor predictor of performance ratings, *23rd Annual Conference of The Society for Industrial and Organizational Psychology*. San Francisco, CA.

Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment, 14*, 317-335.

Blickle, G., Momm, T., Schneider, P. B., Gansen, D., & Kramer, J. (2009). Does acquisitive self-presentation in personality self-ratings enhance validity? Evidence from two experimental field studies. *International Journal of Selection and Assessment, 17*, 142-153.

Borkenau, P., & Ostendorf, F. (1993). *NEO-Fünf-Faktoren Inventar (NEO-FFI) nach Costa und McCrae [NEO-Five-Factor inventory (NEO-FFI) based on Costa and McCrae].* Göttingen, Germany: Hogrefe.

Burns, G. N., & Christiansen, N. D. (2007, April). *Combating the effects of faking: Method factors and discriminate validity.* Paper presented at the 22nd annual meeting of the Society of Industrial and Organizational Psychology, New York.

Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/WINDOWS. Basic concepts, applications, and programming*. Mahwah, NJ: Erlbaum.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: basic concepts, applications, and programming*. Mahwah, NJ: Erlbaum.

Cellar, D. F., Miller, M. L., Doverspike, D. D., & Klawsky, J. D. (1996). Comparison of factor structures and criterion-related validity coefficients for two measures of personality based on the five factor model. *Journal of Applied Psychology, 81*, 694-704.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233-255.

Christiansen, N. D., Burns, G. N., & Montgomery, G. E. (2005). Reconsidering forced-choice item formats for applicant personality assessment. *Human Performance, 18*, 267-307.

Clark, J. M. I., & Biderman, M. D. (2006). A structural equation model measuring faking propensity and faking ability, *21st annual conference of the Society for Industrial and Organizational Psychology*. Dallas, TX.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ, England: Erlbaum.

Collins, J. M., & Gleaves, D. H. (1998). Race, job applicants, and the five-factor model of personality: Implications for black psychology, industrial/organizational psychology, and the five-factor theory. *Journal of Applied Psychology, 83*, 531-544.

Collins, J. M., Schmidt, F. L., Sanchez-Ku, M., Thomas, L., McDaniel, M. A., & Le, H. (2003). Can basic individual differences shed light on the construct meaning of assessment center evaluations? *International Journal of Selection and Assessment, 11*, 17-29.

Conway, J. M., Jako, R. A., & Goodman, D. F. (1995). A meta-analysis of interrater and internal consistency reliability of selection interviews. *Journal of Applied Psychology, 80*, 565-579.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417-440.

Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology, 84*, 155-166.

Ferris, G. R., Perrewé, P. L., & Douglas, C. (2002). Social effectiveness in organizations: Construct validity and research direction. *Journal of Leadership & Organizational Studies, 9*, 49-63.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York, NY: Mcgraw-Hill.

ForsterLee, R. (2007). Personality, interest, and motivational determinants of maximal and

typical performances on a verbal knowledge task. *Human Performance, 20*, 287-292.

Gangestad, S. W., & Snyder, M. (2000). Self-monitoring: Appraisal and reappraisal.

Psychological Bulletin, 126, 530-555.

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure.

Journal of Personality and Social Psychology, 59, 1216-1229.

Griffith, R. L., & Peterson, M. H. (2006). *A closer examination of applicant faking behavior*.

Greenwich, CT: Information Age Publishing.

Guion, R. M. (1965). *Personnel testing*. New York: McGraw-Hill.

Hall, R. J., Snell, A. F., & Foust, M. S. (1999). Item parceling strategies in SEM: Investigating

the subtle effects of unmodeled secondary constructs. *Organizational Research Methods,*
2, 233-256.

Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment

selection. *Journal of Applied Psychology, 92*, 1270-1285.

Hogan, R. (1991). Personality and personality measurement. In M. D. Dunnette & L. M. Hough

(Eds.), *Handbook of industrial and organizational psychology* (pp. 873-919). Palo Alto,

CA: Consulting Psychologists Press.

Hogan, R., Hogan, J., & Roberts, B. W. (1996). Personality measurement and employment

decisions - Questions and answers. *American Psychologist, 51*, 469-477.

Hogan, R., & Shelton, D. (1998). A socioanalytic perspective on job performance. *Human*

Performance, 11, 129-144.

Holden, R. R., Kroner, D. G., Fekken, G. C., & Popham, S. M. (1992). A model of personality

test item response dissimulation. *Journal of Personality and Social Psychology, 63*, 272-

279.

Huffcutt, A. I., Conway, J. M., Roth, P. L., & Stone, N. J. (2001). Identification and meta-analytic assessment of psychological constructs measured in employment interviews. *Journal of Applied Psychology, 86*, 897-913.

Hülsheger, U. R., Maier, G. W., Stumpp, T., & Muck, P. M. (2006). Vergleich kriteriumsbezogener Validitäten verschiedener Intelligenztests zur Vorhersage von Ausbildungserfolg in Deutschland: Ergebnisse einer Metaanalyse [Comparing the criterion-related validities of different intelligence tests for predicting educational success in Germany: A meta-analysis]. *Zeitschrift für Personalpsychologie, 5*, 145-162.

Klehe, U.-C., & Anderson, N. (2007). Working hard and working smart: Motivation and ability during typical and maximum performance. *Journal of Applied Psychology, 92*, 978-992.

Klehe, U.-C., Anderson, N., & Viswesvaran, C. (2007). More than peaks and valleys: Introduction to the special issue on typical and maximum performance. *Human Performance, 20*, 173-178.

Klehe, U.-C., König, C. J., Kleinmann, M., Richter, G. M., & Melchers, K. G. (2008). Transparency in structured interviews-consequences for construct and criterion-related validity. *Human Performance, 21*, 107-137.

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.

Kleinmann, M. (1997a). *Assessment-Center: Stand der Forschung, Konsequenzen für die Praxis [Assessment centers: State of research, consequences for practice]*. Göttingen: Verlag für Angewandte Psychologie.

Kleinmann, M. (1997b). Transparenz der Anforderungsdimensionen: Ein Moderator der Konstrukt- und Kriteriumsvalidität des Assessment-Centers [Transparency of the

requirement dimensions: A moderator of assessment centers' construct and criterion validity]. *Zeitschrift für Arbeits- und Organisationspsychologie, 41*, 171-181.

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.

Komar, S., Brown, D. J., Komar, J., & Robie, C. (2008). Faking and the validity of conscientiousness: A monte carlo investigation. *Journal of Applied Psychology, 93*, 140-154.

König, C. J., Melchers, K. G., Kleinmann, M., Richter, G. M., & Klehe, U.-C. (2006). The relationship between the ability to identify evaluation criteria and integrity test scores. *Psychology Science, 48*, 369-377.

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.

Kuncel, N. R., & Borneman, M. J. (2007). Toward a new method of detecting deliberately faked personality tests: The use of idiosyncratic item responses. *International Journal of Selection and Assessment, 15*, 220-231.

Kunda, Z., & Sanitioso, R. (1989). Motivated changes in the self-concept. *Journal of Experimental Social Psychology, 25*, 272-285.

Latham, G. P., & Wexley, K. N. (1994). *Increasing productivity through performance appraisal* (2nd ed.). Reading, MA: Wesley.

Marcus, B. (2003). Persönlichkeitstests in der Personalauswahl: Sind "sozial erwünschte" Antworten wirklich nicht wünschenswert? [Personality testing in personnel selection: Is

"socially desirable" responding really undesirable?]. *Zeitschrift für Psychologie, 211*, 138-148.

Marcus, B. (2009). 'Faking' from the applicant's perspective: A theory of self presentation in personnel selection settings. *International Journal of Selection and Assessment, 17*, 417-430.

Marcus, B., Goffin, R. D., Johnston, N. G., & Rothstein, M. G. (2007). Personality and cognitive ability as predictors of typical and maximum managerial performance. *Human Performance, 20*, 275-285.

Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology, 35*, 63-78.

McFarland, L. A., & Ryan, A. M. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology, 85*, 812-821.

McFarland, L. A., & Ryan, A. M. (2006). Toward an integrated model of applicant faking behavior. *Journal of Applied Social Psychology, 36*, 979-1016.

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.

Mersman, J. L., & Shultz, K. S. (1998). Individual differences in the ability to fake on personality measures. *Personality and Individual Differences, 24*, 217-227.

Mesmer-Magnus, J., & Viswesvaran, C. (2006). Assessing response distortion in personality tests: A review of research designs and analytic strategies. In R. L. Griffith & M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (pp. 85-114). Greenwich, CT: Information Age Publishing.

Mielke, R., & Kilian, R. (1990). Wenn Teilskalen sich nicht zu dem ergänzen, was die Gesamtskala erfassen soll: Untersuchungen zum Self-Monitoring-Konzept [When subscales don't add up to what the overal construct entails: Research on the concept of self-monitoring]. *Zeitschrift für Sozialpsychologie, 21*, 126-135.

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.

Mischel, W., & Shoda, Y. (1998). Reconciling processing dynamics and personality dispositions. *Annual Review of Psychology, 49*, 229-258.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K. R., & Schmitt, N. (2007a). Are we getting fooled again? Coming to terms with limitations in the use of personality tests for personnel selection. *Personnel Psychology, 60*, 1029-1049.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K. R., & Schmitt, N. (2007b). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology, 60*, 683-729.

Motowidlo, S. J. (1999). Asking about past behavior versus hypothetical behavior. In R. W. Eder & M. M. Harris (Eds.), *The employment interview handbook* (pp. 179-190). Thousand Oaks, CA: Sage.

Motowidlo, S. J., Dunnette, M. D., & Carter, G. W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology, 75*, 640-647.

Mueller-Hanson, R., Heggestad, E. D., & Thornton, G. C. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. *Journal of Applied Psychology, 88*, 348-355.

Nowack, W., & Kammer, D. (1987). Self-presentation: Social skills and inconsistency as

    independent facets of self-monitoring. *European Journal of Personality, 1*, 61-77.

Ones, D. S., Dilchert, S., Viswesvaran, C., & Judge, T. A. (2007). In support of personality

    assessment in organizational settings. *Personnel Psychology, 60*, 995-1027.

Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality

    testing for personnel selection: The red herring. *Journal of Applied Psychology, 81*, 660-

    679.

Pauls, C. A., & Crost, N. W. (2005). Effects of different instructional sets on the construct

    validity of the NEO-PI-R. *Personality and Individual Differences, 39*, 297-308.

Preckel, D., & Schüpbach, H. (2005). Zusammenhänge zwischen rezeptiver

    Selbstdarstellungskomopetenz und Leistung im Assessment Center [Correlations between

    receptive self-presentation competence and performance in an assessment center].

    Zeitschrift für Personalpsychologie, 4, 151-158.

Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The impact of response

    distortion on preemployment personality testing and hiring decisions. *Journal of Applied

    Psychology, 83*, 634-644.

Sackett, P. R. (2007). Revisiting the origins of the typical-maximum performance distinction.

    Human Performance, 20, 179-185.

Sanitioso, R., Kunda, Z., & Fong, G. T. (1990). Motivated Recruitment of Autobiographical

    Memories. *Journal of Personality and Social Psychology, 59*, 229-241.

Schmit, M. J., & Ryan, A. M. (1993). The Big-5 in personnel-selection: Factor structure in

    applicant and nonapplicant populations. *Journal of Applied Psychology, 78*, 966-974.

Schmitt, N., & Oswald, F. L. (2006). The impact of corrections for faking on the validity of

    noncognitive measures in selection settings. *Journal of Applied Psychology, 91*, 613-621.

Schollaert, E., & Lievens, F. (2008). The effects of exercise instructions on the observability of assessment center behavior. *International Journal of Psychology, 43*, 577-577.

Smith, D. B., Hanges, P. J., & Dickson, M. W. (2001). Personnel selection and the five-factor model: Reexamining the effects of applicant's frame of reference. *Journal of Applied Psychology, 86*, 304-315.

Smith, M. A., Moriatry, K. O., Lutrick, E. C., & Canger, J. M. (2001). *Exploratory factor analysis of NEO PI-R for job applicants: Evidence against the Big Five.* Paper presented at the 16th Annual Meeting of the Society for Industrial and Organizational Psychology, San Diego, CA.

Snell, A. F., Sydell, E. J., & Lueke, S. B. (1999). Towards a theory of applicant faking: Integrating studies of deception. *Human Resource Management Review, 9*, 219-242.

Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology, 30*, 526–537.

Stark, S., Chernyshenko, O. S., Chan, K. Y., Lee, W. C., & Drasgow, F. (2001). Effects of the testing situation on item responding: Cause for concern. *Journal of Applied Psychology, 86*, 943-953.

Stricker, L. J. (1969). "Test-wiseness" on personality scales. *Journal of Applied Psychology, 53*, 1-17.

Tett, R. P., & Christiansen, N. D. (2007). Personality tests at the crossroads: A response to Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt (2007). *Personnel Psychology, 60*, 967-993.

Topping, G. D., & O'Gorman, J. G. (1997). Effects of faking set on validity of the NEO-FFI. *Personality and Individual Differences, 23*, 117-124.

Van Iddekinge, C. H., Raymark, P. H., Eidson, C. E., & Putka, D. J. (2001). *Applicant-incumbent differences on personality, integrity, and customer service measures.* Paper presented at the 16th Annual Meeting of the Society for Industrial and Organizational Psychology, San Diego, CA.

Vasilopoulos, N. L., Cucina, J. M., Dyomina, N. V., Morewitz, C. L., & Reilly, R. R. (2006). Forced-choice personality tests: A measure of personality and cognitive ability? *Human Performance, 19*, 175-199.

Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement, 59*, 197-210.

Whyte, W. H., Jr. (1956). *The organization man*. New York: Simon & Schuster / Touchstone.

Wiggins, J. S. (1959). Interrelationships among MMPI measures of dissimulation under standard and social desirability instructions. *Journal of Consulting Psychology, 23*, 419-427.

Woehr, D. J., & Huffcutt, A. I. (1994). Rater training for performance appraisal: A quantitative review. *Journal of Occupational and Organizational Psychology, 67*, 189-205.

Wrensen, L. B., & Biderman, M. D. (2005). Factors related to faking ability: A structural equation model application, *20th annual conference of the Society for Industrial and Organizational Psychology*. Los Angeles, CA.

Zickar, M. J., & Robie, C. (1999). Modeling faking good on personality items: An item-level analysis. *Journal of Applied Psychology, 84*, 551-563.

Table 1.

Means, Standard Deviations, Internal Consistencies, and Correlations Between the Study Variables.

	M	*SD*	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
Performance												
1. Performance high-fidelity simulations	2.22	0.42	(.76)									
2. Performance low-fidelity simulations	2.50	0.45	.47**	(.77)								
Ability to identify criteria (ATIC)												
3. ATIC	1.76	0.56	.27**	.26**	(.84)							
Big Five Personality Dimensions												
4. Openness to experience	3.79	0.45	.12	.19*	.17*	(.72)						
5. Conscientiousness	4.14	0.44	.15	.29**	.30**	.13	(.79)					
6. Extraversion	3.79	0.52	.22**	.14	.12	.22**	.10	(.80)				
7. Agreeableness	3.90	0.43	.17*	.22**	.26**	.22**	.30**	.27**	(.72)			
8. Emotional stability	3.79	0.64	.22**	.28**	.24**	.11	.35**	.51**	.41**	(.88)		
Control variables												
9. Self-monitoring	0.75	0.20	.05	.08	.08	.06	-.06	-.08	.19*	-.23**	(.65)	
10. General mental ability	11.91	2.03	.38**	24**	.17*	.13	.09	.20*	.06	.17*	-.05	(.63)

Note. *N* = 149; * *p* < .05 (two-tailed). ** *p* < .01 (two-tailed).

Table 2.Goodness-of-Fit Indices and Model Comparisons for the Structural Equation Models tested

Model	χ^2	df	p	χ^2/df	TLI	IFI	CFI	RMSEA	RMSEA upper 90%	Model comparison	$\Delta\chi^2$	df	P
Hypothesis 1: Factor Structure of Measurement Models													
1: Five factors	161.29	80	<.01	2.01	.89	.92	.92	.08	.10				
2: Five factors plus IEF	96.97	65	<.01	1.49	.95	.97	.97	.06	.08	1 – 2	64.32	15	<.01
Hypothesis 2: High factor loadings on Conscientiousness and Emotional Stability													
Models proposed to render a good fit, factor loadings proposed to be equal to the highest factor loading													
3a: ES1 = C1	97.01	66	.01	1.47	.95	.97	.97	.06	.08	3a – 2	.04	1	.84
3b: ES1 = ES2	97.35	66	.01	1.48	.95	.97	.97	.06	.08	3b – 2	.38	1	.54
3c: ES1 = C2	97.89	66	.01	1.48	.95	.97	.97	.06	.08	3c – 2	.92	1	.34
3d: ES1 = C3	97.64	66	.01	1.48	.95	.97	.97	.06	.08	3d – 2	.67	1	.41
3e: ES1 = ES3	101.86	66	<.01	1.54	.94	.97	.96	.06	.08	3e – 2	4.89	1	.03
Models proposed to render a poorer fit, factor-loadings proposed to be significantly smaller than the highest factor loading													
3f: ES1 = E2	101.18	66	<.01	1.53	.94	.97	.96	.06	.08	3f – 2	4.21	1	.04
3g: ES1 = E3	103.27	66	<.01	1.56	.94	.96	.96	.06	.08	3g – 2	6.30	1	.01
3h: ES1 = E1	110.90	66	<.01	1.68	.93	.96	.95	.07	.09	3h – 2	13.93	1	<.01
3i: ES1 = A2	101.73	66	<.01	1.54	.94	.97	.96	.06	.08	3i – 2	4.76	1	.03
3j: ES1 = A1	105.34	66	<.01	1.60	.94	.96	.96	.06	.09	3j – 2	8.37	1	<.01
3k: ES1 = A3	125.99	66	<.01	1.91	.90	.94	.94	.08	.10	3k – 2	29.02	1	<.01
3l: ES1 = O1	100.37	66	<.01	1.52	.94	.97	.97	.06	.08	3l – 2	3.40	1	.07
3m: ES1 = O2	101.75	66	<.01	1.54	.94	.97	.96	.06	.08	3m – 2	4.78	1	.02
3n: ES1 = O3	104.35	66	<.01	1.58	.94	.96	.96	.06	.09	3n – 2	7.38	1	<.01

Note.   N = 149. IEF = ideal-employee factor; ES = emotional stability; C = conscientiousness; A = agreeableness; O = openness to experience; E = extraversion; TLI / IFI / CFI = Tucker-Lewis / Incremental / Comparative Fit Index; RMSEA = root mean square error of approximation.

U

Table 3.

Hypothesis 3. ATIC accounting for the criterion related validity of the ideal-employee factor for predicting performance.

Model	χ^2	df	p	χ^2/df	TLI	IFI	CFI	RMSEA	RMSEA upper 90%	Model comparison	$\Delta\chi^2$	df	p
4a: Performance predicted by ATIC, IEF, Big Five, and control variables. ATIC linking the IEF and performance	331.59	231	<.01	1.44	.91	.93	.93	.05	.07				
4b: Model 4a without the links from ATIC to the IEF and performance	359.42	233	<.01	1.54	.88	.91	.91	.06	.07	5b – 5a	27.83	2	< .01

Note. N = 149. ATIC = ability to identify criteria; IEF = latent ideal-employee factor; TLI = Tucker-Lewis Index; IFI = Incremental Fit Index; CFI = comparative fit index; RMSEA = root mean square error of approximation.
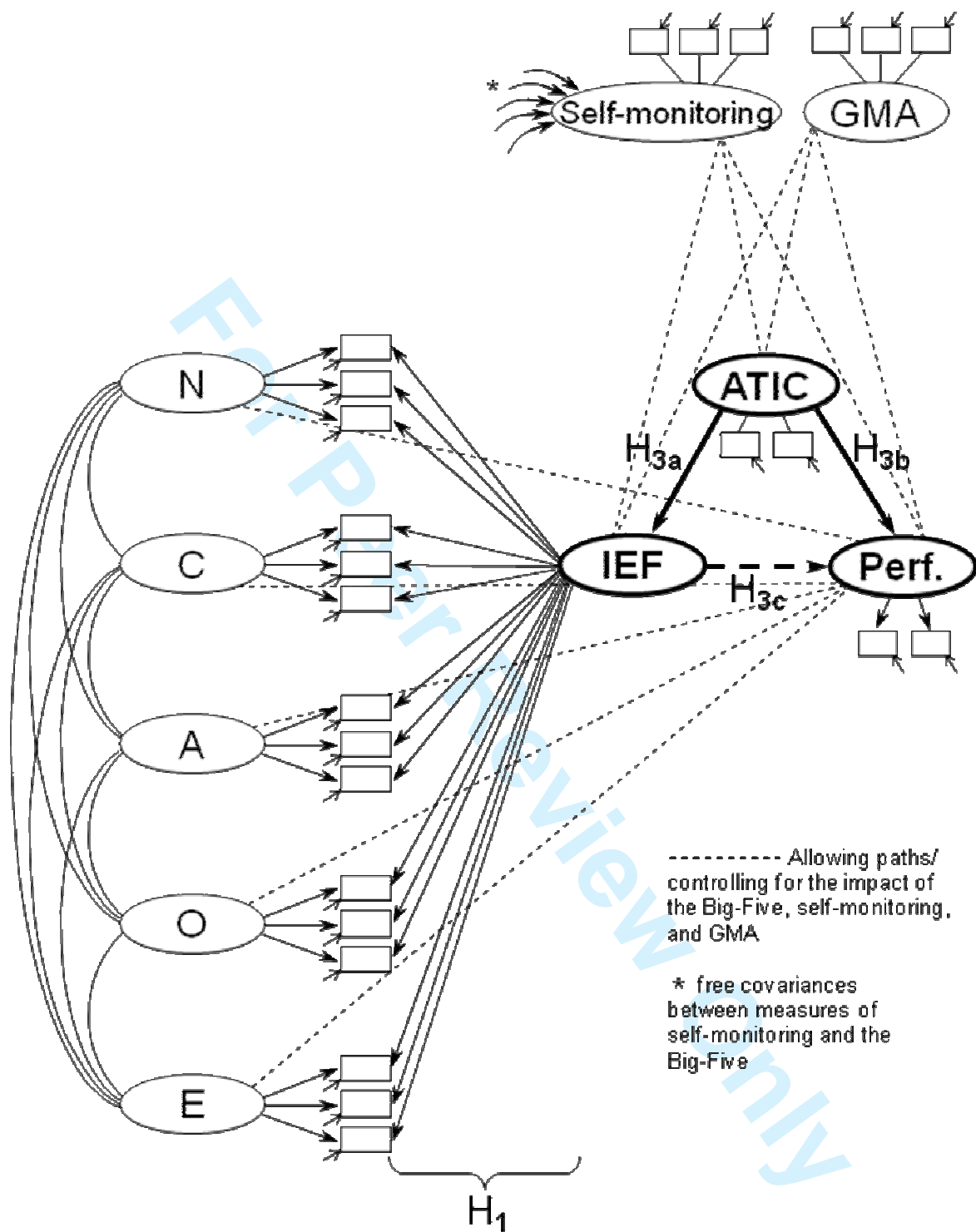
Figure 1.

Proposed model.

The Big Five personality scales load on the same ideal-employee factor (IEF; Hypothesis 1). This

IE-factor predicts performance in diverse work-simulations (Hypothesis 3c), also after accounting

for self-monitoring and general mental ability (GMA). Participants' ability to identify criteria

(ATIC) relates to both the IEF (Hypothesis 3a) and performance (Hypothesis 3b), fully

accounting for the relationship between the IE-Factor and performance (Hypothesis 3c). (ES =

emotional stability, C = conscientiousness, A = agreeableness, O = openness to experience, E =

extraversion, low/high fid = low/high fidelity simulation; dashed paths = controlling for self-

monitoring and GMA).
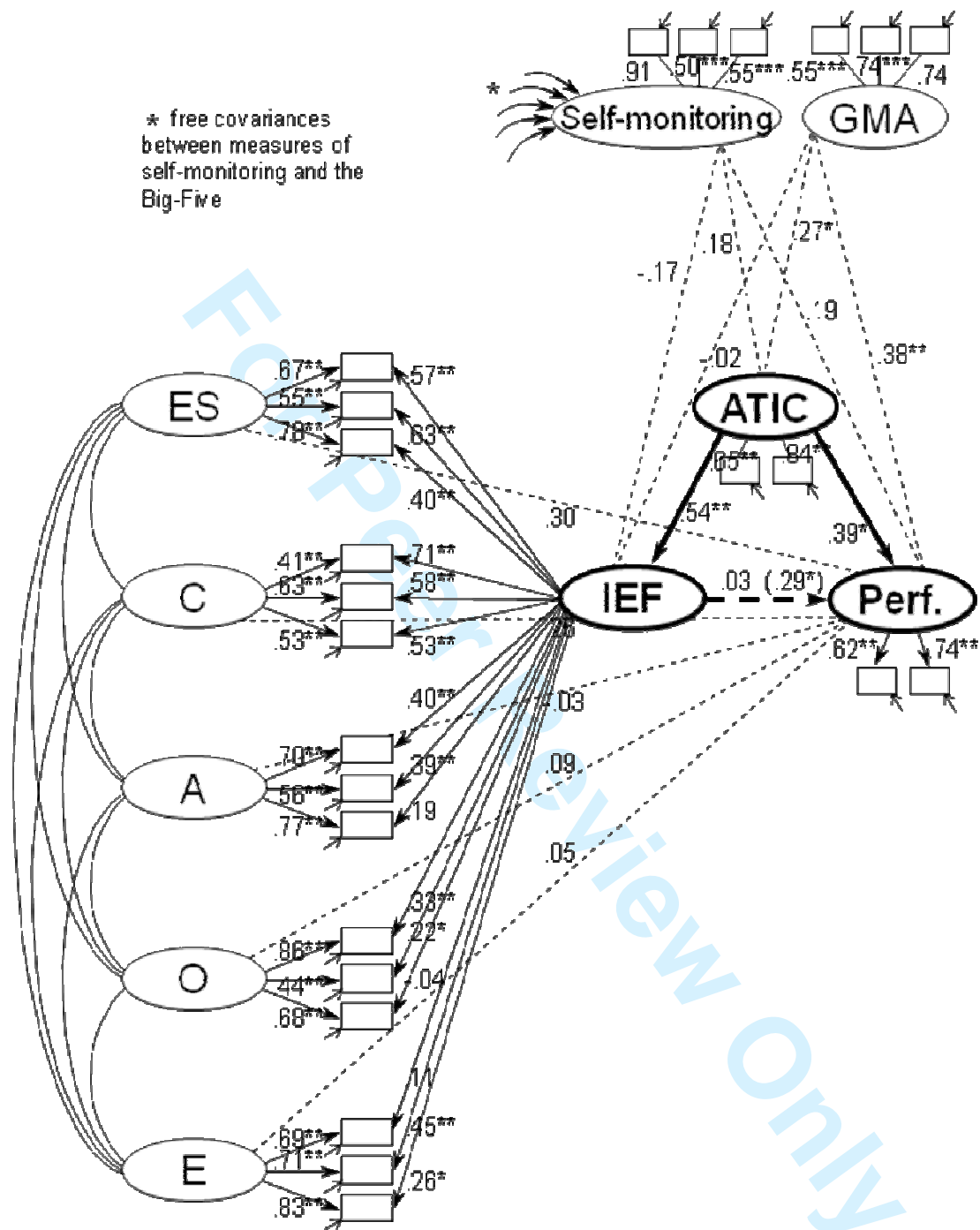
Figure 2.

Final model.

Significant paths (p < .05; two-tailed) are depicted in bold, non-significant paths are dashed. (ES

= emotional stability, C = conscientiousness, A = agreeableness, O = openness to experience, E =

extraversion, IE = ideal-employee factor, ATIC = ability to identify criteria, GMA = general

mental ability, Perf. = performance, low/high fid = low/high fidelity simulation). The value in

parentheses describes the weight from the IEF to performance before the inclusion of the two

paths leading from ATIC to both the IEF and to performance.

Figure 1.

---------- Allowing paths/
controlling for the impact of
the Big-Five, self-monitoring,
and GMA

* free covariances
between measures of
self-monitoring and the
Big-Five

Figure 2.

* free covariances
between measures of
self-monitoring and the
Big-Five

Appendix A: *Job advertisement*

We are looking for applicants from all fields of study for our

Trainee Program

With more than five thousand employees and a transaction volume of about €500 million in 2005, RETEC is one of the biggest diversified technology companies in Germany, serving customers and communities with innovative products and services in each of our businesses. As dynamic as the development of our business is the increasing number of our employees.

Are you looking for new challenges in a growing company?

We're committed to hiring new university graduates from various educational backgrounds for our trainee-program *protec*. This 12 to 18 month program offers challenging and developmental project assignments. You will learn about diverse operations from our broad field of task resorts located throughout Germany. For example, you might learn about marketing- or personnel management. A refined qualification program with individualized personnel development plans as well as diverse activities organized for our young talent groups within the company will further help you to build social networks and to develop professionally as well as personally.

You like our program? If you are someone who can deal with ever-growing complexity, who thrives during teamwork, and if you are poised to accept responsibility for yourself and others, we are looking forward to your meaningful application documents.

Appendix B

Pedro Brandao about the trainee-program *protec*

"I have been working in the protec-team for five months now and am truly enthused by the variety of possibilities the program offers. My first assignment was at the head office in Frankfurt, where I was assigned a highly topical project within the merger between CAP SOTIE and retec: I was asked to develop a common concept for the sales- and distribution data base and to present this concept to management. Now I've accepted a three-month project in Lüneburg in the field of corporate HR-development.

I can just congratulate everyone who finds their way into the program. From the beginning, I've been fully integrated into the operating process and have found enormous support. Also the community within the program and the assistance by others is very good – a crucial ingredient for success in this program, I think. One continuously accepts responsibility for the smooth process of challenging projects as well as for the participating employees – not always an easy task at that degree of complexity! And our daily work is everything *but routine, that's what I appreciate about working with retec."*

Endnotes.

1.   An example: A participant assumes that in the first group discussion participants were evaluated on "teamwork", "creativity", and "goal setting". In the second group discussion, this participant assumed that she was evaluated on "teamwork" and "influencing others". Later, she assumed for the first discussion that the hypothesis "teamwork" fit completely (rating = 4) with the dimension "Cooperation", that "creativity" somewhat (rating = 1) reflected "Handling of Information", and that "goal setting" rather well (rating = 3) reflected "Planning". For the second group discussion, she rated the strength of fit between "teamwork" and "Cooperation" again with 4 and the strength of fit between "influencing others" and "Leadership" with a 4 as well.

In summary, this participant thus received an ATIC score of 4 for Cooperation in both group discussions. In the first group discussion, she also received an ATIC score of 3 for Planning, but a score of 0 for Leadership. In the second group discussion, she received a score of 4 for Leadership and score of 0 for Planning. In average, this would imply an overall ATIC score of 2.5 (ATIC = (4+4+3+0+0+4)/6).

2.   Besides being a statistical necessity in order to render an identifiable solution (Byrne, 1994, 1998; see also Cellar et al., 1996, p. 699), the absence of covariances between the original personality dimensions and the additional ideal-employee factor is warranted for both conceptual and empirical reasons. Conceptually, the ideal-employee factor results from a cognitive schema associated with the job application situation (Fiske & Taylor, 1991; Holden et al., 1992; Mischel & Shoda, 1995, 1998), rather than representing another aspect of personality (Cellar et al., 1996; Van Iddekinge et al., 2001). Empirically, this assumption has been supported by different studies using different methodological approaches that found that faking effects were independent of person effects (Pauls & Crost, 2005) and that the increased common variance seems unrelated to the personality test content variance (Zickar & Robie, 1999).

3. As would have been expected, a model-comparison based on the 60 individual items rather than the fifteen parcels yielded a miserable fit for both Models 1 and 2 on the usual goodness to fit measures (IFI, TLI, and CFI = .54 to .57 for Model 1 and .61 to .65 for Model 2). Yet, this comparison, too, confirmed that Model 2, $\chi^2(1640) = 2787.01$, $p < .01$, $\chi^2/_{df} = 1.70$, assuming all 60 individual items to load on the ideal-employee factor, yielded a significantly better fit than the baseline measurement Model 1, $\chi^2(1700) = 3085.68$, $p < .01$, $\chi^2/_{df} = 1.82$, $\Delta\chi^2(60) = 298.67$, $p < .01$. This, again, confirms the necessity of including an ideal-employee factor to the model and shows that results mentioned above are not caused by any distribution of items onto parcels. The average loading of the individual items onto the IEF was $\gamma = .46$ for the items belonging to Emotional Stability, $\gamma = .39$ for the items belonging to Conscientiousness, $\gamma = .18$ for the Extraversion items, $\gamma = .14$ for the Agreeableness items, and $\gamma = .08$ for the Openness items, respectively.

4. We gained these estimates for each personality item's loading onto the ideal-employee factor via the item-based measurement Model 2 mentioned in footnote 3.