Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht.

Mitglieder der Prüfungskommission:

Vorsitz: Prof. Dr. Dr. Jürgen Henning, Dekan, Fachbereich 06

Erstgutachter: Prof. Dr. Ute-Christine Klehe
Arbeits- und Organisationspsychologie
FB 06, JLU Gießen

Zweitgutachter: Prof. Dr. Martin Kersting
Psychologische Diagnostik
FB 06, JLU Gießen

Weitere Professoren: Prof. Dr. Frank Walter
Betriebswirtschaftslehre VII
FB 02, JLU Gießen
Prof. Dr. Barbara Holland-Cunz
Gender Studies
FB 03, JLU Gießen

**Acknowledgment**

To all the people who contributed to this dissertation- THANK YOU!

Ute-Christine Klehe, thank you for seeing my talent, drive and ambition when I was an undergraduate student and believing in me as a (junior) scientist at the time. Thank you for your supervision and advice concerning my research, and my dissertation.

Thank you to all my PhD fellows. Some of you showed me how it is done and lead by example. All the nights brainstorming, complaining and motivating one another, helped me to get to this stage. Most of all, I would like to thank my colleague Maximillian Buyken. For a few years, you were my office buddy and became a good friend. I not only valued your knowledge and advice, but also your points of view in stressful situations and the motivation "to stick with it". Now it is your turn!

Thank you to my life partner Visvesh Seshadri. You followed me to Germany for something that would show you a different me. Yet, you are still by my side loving and supporting me with everything I do. You gave and still give me a home, where everything is possible.

Hanaleï, you made me realize how hard small things can be to learn but how attainable things can be when you have everything you need to accomplish something great. You motivated me to sit down and do the work.

Last but not least, Marcel: Even though you lost the hardest battle in life, you were the strongest and bravest I have ever seen you. You tried to keep all the sorrows away from us but clearly showed us that some things just need to be done and with others you don't need to waste time. When I was complaining about how much more I have to do in order to finish my dissertation and doubting whether I can, you looked straight at me and said: Vanessa, du machst das Ding jetzt fertig! I did!

## Table of Contents

## List of Figures

# List of Tables

# General Introduction

## When to Show Who You Are or Who You Could Be:

## Personnel Selection and Finding the Right Person for the Job

"If you can't make it good, at least make it look good."

Bill Gates

Most job applicants around the world want "to make it good," and if this is not possible because they lack the right education, qualifications, skills, abilities, or other job requirements, some of them may try to "at least make it look good."

Organizations looking for a new employee have the goal to distinguish between the applicants who are good and applicants who only seem good. An organization's aim is to identify applicants who are high performers, good and trustworthy employees, and, in short, have the potential to be the perfect employee. Job seekers and employers have a few motivations in common. The former want to be right for the job, get a job offer, and be happy at work. The latter want the right person for the job, make a job offer, and have a happy employee. However, job seekers and employers may use different means to get what they want.

The goal of personnel selection is to fairly and objectively identify the right candidate for the job, i.e., the applicant who matches the job requirements and is expected to achieve high work performance (Schmidt & Hunter, 1998). To achieve that goal, organizations on the one hand implement a variety of selection tools to maximize selection success. On the other hand, applicants strive to make the right impression, i.e., to present themselves as a perfect match, even if they lack some of the requirements (Bozeman & Kacmar, 1997). Nevertheless, the overall purpose of utilizing selection tools is to differentiate those who are the right candidate from those who may only seem to be the right candidate.

The current line of research studies whether (some) personnel selection tools foster such a differentiation, or actually enable applicants to present themselves more favorably than they actually are. Further, it explores why some applicants present themselves more favorably than others, and when and how they engage in such impression management (IM). Lastly, the current research illustrates how some selection methods intent to be fair, but with that intention actually disadvantage some applicants by causing them to perform worse, and others to perform better, i.e., even though some could "make it good", the selection tool prevents applicants from "making it look good."

## Personnel Selection

Organizations often rely on different selection tools to identify a successful employee (Schmidt & Hunter, 1998). Most often, the first initial step to assess candidates' suitability is with the help of biographical data, e.g., cover letter and curriculum vitae (CV). From an employer's view, the goal is to determine whether candidates possess the needed experience, skills and abilities to perform the job successfully (Knouse, 1994). Thus, the organizations rely on data written by candidates for a first impression. With the information candidates provide, managers try to predict future work performance and decide who should be granted a chance of proving his or her suitability in the next selection step, often in the form of a (structured) interview or written test.

From a candidate's perspective, this is the first chance of making the "right" impression, i.e., trying to look as good as possible, impressing the potential new employer, and maximizing ones chance of being invited to the next selection step. Schmidt and Hunter (1998) found that biographical data as well as past job experiences predict future job performance. Similarly, Rothstein, Schmidt, Erwin, Owens, and Sparks (1990) showed that the validity of biographical data could be generalized across jobs, positions, and time. Thus, research agrees that biographical data is indeed a valid selection tool. However, it is possible that candidates

will use his/her CV and cover letter, therewith the provision of his/her personal biographical data, to enhance picture of him-/herself to increase his or her chances of getting the job. Nevertheless, even though CVs and cover letters may have some suggested formats and expected information, the candidate is free to write anything he or she wants in a way he or she desires.

Indeed, Knouse (1994) found that candidates who engaged in some form of IM were perceived as more confident and more hirable. In a similar study, Knouse, Giacalone, and Hinda (1988) concluded that IM could also decrease the perceived likability of the candidate. Even though these studies illustrated that IM in cover letters and CVs can influence the decision of which candidate would be invited to an interview, research that tackles IM on biographical data is rather rare. Only one recent study has examined whether candidates actually engage in IM in their cover letters and CVs and if so, which tactics they employ. No study so far has examined whether candidates who employ IM on cover letter and CV are more likely to continue with their IM behaviours during a selection interview, compared to candidates who do not show IM behaviours on their cover letter and CV. Therefore, this first study in this line of research aims to examine the IM behaviors of candidates in this initial first contact with the employer and whether IM behavior during this step influence whether candidates proceed to again engage in IM in the interview or testing phase. Many studies have investigated IM during a job interview and clearly concluded that IM does influence interview performance and in turn hiring decisions (e.g., Barrick & Mount, 1996; Ellis, West, Ryan, & DeShon, 2002b; Kleinmann & Klehe, 2010). However, no study to date has examined whether IM behavior is a stable individual disposition or whether it depends on the situation. In other words, this first study further explores whether candidates' IM in his/her cover letter and CV predict IM during an interview, or whether the situation (cover letter/CV vs. interview) determines the usage of IM.

IM plays a crucial role also in other selection tools; often, organizations also rely on personality tests to make job performance predictions. Past research has shown that these tests are also influenced by candidates' IM, often labeled as response distortion (e.g. Rosse, Stecher, Miller, & Levin , 1998). Compared with a normal or honest situation when candidates answer personality items to represent their true self, candidates understandably are inclined to answer personality items in a way that makes them "look good" when in a selection situation. Past research has shown that candidates indeed distort their answers, scoring particularly high on conscientiousness and emotional stability, which are likely important for all jobs (Schmit & Ryan, 1993; Bikeland et al. 2006). Thus, whether candidates see themselves as potentially very conscientious or emotionally stable or whether they only want to make the impression, past research has indicated that candidates want to display the image of an ideal employee. Yet more than a mere score inflation, the factor structure of the inventory suffers too. The original 5 factor structure that the inventory had been designed for, does not hold true under selection circumstances. In selection circumstances a 6-factor structure appears to represents the personality constructs better (Schmit & Ryan, 1993). An additional factor composed of loadings of all but mostly of two (conscientiousness and emotional stability) personality traits emerged, which Schmit and Ryan (1993) labeled as the ideal-employee factor (IEF). Although the majority of research has focused on the Big Five personality traits under personnel selection, no study has addressed if and how the pattern of the IEF changes depending on the job description. Relying on the person-situation perspective (Fiske & Taylor, 1991; Mischel & Shoda, 1995, 1998), we argue that the pattern of the IEF changes according to the job profile, i.e., personality traits that are more relevant to the job will load higher onto the IEF. Additionally, this second study extends the current knowledge on the IEF by illustrating that response distortion causes both score inflation and the emergence of an additional factor. Past studies often either lacked the non-evaluative (honest) condition to compare with the

evaluative (selection) condition. With samples from both conditions, we were able to examine whether score inflation is indeed the same as the emergence of the IEF. However, the candidates may change their performance consciously or unconsciously and may seem to be someone they are not. Additionally, selection tools, or the way the selection tools are presented, may unintentionally influence candidates' performance, i.e., the candidates' shown behavior does not actually represent their level of skills or abilities.

Transparency, i.e., providing information on what the selection tool actually measures, has been widely accepted to increase the fairness of personnel selection. The underlying argument is that the knowledge of what the test or exercise measures, helps all candidates to focus on relevant behavior, implement behavior that is part of the performance criteria, and in turn increase performance (e.g., Kleinmann, 1993; Kleinmann, Kuptsch, & Köller, 1996). Indeed, past studies have shown that such transparent selection exercises increased performance and perceived fairness by candidates. Candidates then do not have to wonder what the exercise aims to assess, but instead can focus on implementing the desired behavior.

However, the third and last study of this dissertation shows that the relationship between transparency and performance is not so straightforward. In fact, it depends on the candidate and the information revealed as to whether transparency leads to a performance boost. Relying on the stereotype threat theory (Steele & Aronson, 1995) and social identity theory, the third study shows that performance only increases for those whose social identity is not threatened by the information presented. The underlying argument is that if candidates know that a selection exercise aims to assess a certain dimension that holds a negative stereotype about them, those candidates may experience fear to confirm the negative stereotype, whereas other candidates experience a boost and perform better because the same dimension reveals a positive stereotype about them. With two different selection tools, we illustrate that transpar-

ency does not lead to a performance increase for all candidates, only for some. Thus, the intention to make selection tools transparent is a noble one, but it might not enable some candidates to show their full potential.

Overall, this line of research aims to studies how the impression and performance of candidates can be influenced, either in relation to the instrument or by the candidate. The following three studies not only move chronologically from the first selection step (cover letter and CV and interview) through a written test (Big Five Personality Test) to a final selection step (assessment center exercise), but also aims to illustrate strengths and weaknesses of each selection step.

# Chapter 1:

# I am Great!

# What Candidates State or do not State: Impression Management Throughout the Selection Process

Vanessa Jacksch & Ute-Christine Klehe

Justus-Liebig-University Giessen

Nicolas Roulin

University of Manitoba

Author Note:

Vanessa Jacksch & Ute-Christine Klehe, Department of Work & Organizational Psychology, Justus-Liebig-University

Nicolas Roulin, Department of Business Administration, University of Manitoba

## Abstract

Impression management (IM) is a constant concern during personnel selection. For one, during an interview or assessment exercise and such IM can influence decision making (e.g., Barrick & Mount, 1996; Ellis, West, Ryan, & DeShon, 2002a). First impression count, thus IM behaviors can influence the impression an candidate leaves even during the first contact point between candidate and employer, yet no study to date has explored whether candidates already engage in IM behaviors during the first step of their application, i.e., in their cover letters and curriculum vitaes (CV). The current study coded cover letters and CVs on IM tactics and shows that candidates indeed engage in several IM tactics, and further illustrates which IM tactics are used and when. Additionally, the study examines if some people tend to use IM throughout the selection process or whether it depends on the assessment situation, as well as which individual antecedents may influence IM behavior.

*Keywords:* impression management, cover letter, curriculum vitae, interview

# I am Great!

## What Candidates State or do not State:

## Impression Management Throughout the Selection Process.

"Whenever you are asked if you can do a job, tell'em, 'Certainly I can!' Then get busy and find out how to do it."

Theodore Roosevelt

Already Theodore Roosevelt advised people to represent themselves the best way possible when applying for a job. Over the past decades, many studies have focused on how people attempt to control the image that people form about oneself. Such an attempt is often referred to as impression management (IM) and a vast variety of studies have focused on IM during personnel selection (Barrick & Mount, 1996; Bolino, Kacmar, Turnley, & Gilstrap, 2008; Ellis, West, Ryan, & DeShon, 2002b; Kleinmann & Klehe, 2010; Leary & Kowalski, 1990; McFarland, Ryan, & Kriska, 2003; Peeters & Lievens, 2006). However, even though the number of studies researching IM during personnel selection is high, most studies have focused on IM in the context of job interviews. But what happens before? Researchers paid much less attention on how IM may influence impressions at the very first stage of selection, i.e. curriculum vitae (CV) and cover letter. The very first impression an candidate is able to make is with his or her written application for the job, but very few studies have researches IM during this particular selection step (Knouse, 1994; Knouse, Giacalone, & Hinda, 1988). Additionally, those studies only focused on the effects that IM in a cover letter or CV might have on the person who evaluates the candidates' documents. To our knowledge, only one very recently published study has focused on whether job candidates actually engage in IM in their cover letter and CV and what IM tactics are being used (Waung, McAusian, DiMambro and Miegoc, 2017) and no study so far has also explored whether candidates who use IM in

their written documents do so in interviews as well. The need for more knowledge about IM at the stage of CV and cover letter is high, as organizations use them in high numbers to screen a large number of job candidates. Past research shows that IM during a job interview can influence hiring decision (Bolino, Kacmar, Turnley, & Gilstrap, 2008; Gilmore & Ferris, 1989) and the few exploring the effects of IM on cover letter and CV show that it also influence the impression an candidate leaves.

The current study therefore aims to systematically examine IM tactics use in cover letter and resume and therewith tries to answer the following questions. First, do candidates show IM tactics in their cover letter and CV and how much IM behaviors are shown? Second, what IM tactics are used? Third, does the implementation of IM depend on the situation the candidate is in, or does it depend on the candidate? Therewith we hope to make the following contributions. First, we like to increase the awareness among researchers as well as practitioners that IM already plays a role on cover letter and CV. Decisions about who is invited to an interview may already be affected by the IM behaviors of candidates during the initial first contact to the employer. Second, by studying whether IM depends on the person or the situation, we hope to find ways to anticipate IM behavior and if such behavior is not desired, to find tools to minimize it. Third, with identify more antecedents of IM and what IM behaviors are particularly often shown may help us and practitioners understand how we can change our personnel selection tools to either counter-balance, control or embrace IM behaviors during personnel selection.

For now, what a candidate has to do is to impress. Namely, whether a candidate is invited to a job interview mostly depends on the impression he or she leaves with his or her cover letter and CV. Thousands of cover letters and CVs are send to organizations world-wide (Bradford, 2012). In order to be even invited to a job interview, those candidate materials need to impress positively. As past studies have indicated, IM can influence whether a person

is rated positively or negatively during a job interview (Barrick & Mount, 1996; Roulin & Krings, 2016). It is therefore possible that the same is true for cover letters and CV. The following study therefore aims to shed light on IM through several selection steps, starting with cover letter and CV

**IM in Cover Letters, CVs**

Knouse (1994) argued that IM in cover letters and CVs could influence the perception practitioner might have of a candidate. Candidates who seemingly engaged in IM scored higher on their interpersonal skills, self-confidence, and hireability beyond education or work experience. In contrast, Knouse's earlier study in 1988 found that IM actually influenced ratings in a mostly negatively manner. Candidates who seemingly engaged in IM were perceived as less likable, less truthful, and less employable but higher in self-confidence compared with candidates who appeared not to have engaged in IM. Since the focus of both studies was on the effect of IM in cover letters and CVs, however, Knouse used 'faked' cover letters and CVs to accurately compare cover letters and CVs, so that educational information, work experience, etc. were held constant and only the amount of IM varied. Since IM in cover letters and CVs does influence readers' perceptions of a candidate beyond the candidate's actual qualifications, this increases the importance of studying IM during that first initial contact with the employer. During interviews, almost all job candidates use IM (e.g. Ellis, West, Ryan, & DeShon, 2002b; Griffith, Chmielowski, & Yoshita, 2007; Levashina & Campion, 2006; Stevens & Kristof, 1995), thus likely that almost all candidates will already do so on their cover letters and CVs, as the underlying motivation during both situation is the same: make a favourable impression and get the job (interview). The motivation to impress pushes candidates to show IM behaviors.

Thus, our first focus is to explore whether IM is used and what tactics are employed.

We propose:

*Hypothesis 1:* Job candidates engage in IM behaviors in their cover letters and CVs.

## IM Tactics Depending on the Selection Method

Candidates may act as actors throughout the selection process, though their performance and behavior may depend on the characteristics of both the situation and the audience (Goffman, 1959). For example, Ellis et al. (2002) showed that the interview format determined to what extent candidates engaged in different IM tactics. Further, IM was less prevalent during structured interviews than unstructured interviews. Similarly, McFarland et al. (2003) found that IM is used less frequently during a role-play compared with an interview.

Additionally IM tactics may differ depending on the selection situation. During job interviews, candidates tend to use more assertive tactics of self-promotion and ingratiation (e.g. Stevens and Kristof, 1995). Self-promotion tactics aim to enhance the impression of competence and accomplishment by using strategies such as providing examples of successful past behavior (exemplification), taking responsibility for positive events (entitlement) or overstating the success of value of a particular event (enhancement) or simply describing oneself and past (work) experience positively. Contrastingly, ingratiation tactics are focused at enhancing likeability, including strategies such as other enhancement (paying a compliment) and opinion conformity.

The cybernetic model of IM suggests that IM tactics that are expressed depend on the situation and the interaction partner (Bozeman & Kacmar, 1997). In other words, IM behaviors are aimfully used, directed at a target. A person who uses IM directs that behavior at another person who judges whether the person is likable, confident and suited for the job. During personnel selection, such a target could be the interviewer during a selection interview or a target could also be the impression or image a candidate has about an organization he or she

wants to work for. With that image in mind, a candidate might write a job application according to that held image. Similarly, during an interview, the interviewee tries to establish an impression that is consistent with the competences required on the job, that he or she fits in well and is liked. Thus depending on the target, candidates might use different IM tactics. During a job interview the interviewer is the IM's target. If a chosen tactic leads to success, i.e. the interviewee has the impression that the interviewer approves of his/her statement or answer, the interviewee might continue to reinforce the created image by using the same IM tactics (Bozeman & Kacmar, 1997). If the IM tactic does not lead to success, however, the candidate might try other IM tactics to test their effectiveness, hoping the other IM tactic will cause a positive impression. Thus, during an interview, the candidate is able to use the interaction partner as a feedback provider, creating the possibility to try out several different IM tactic. Due to a lack of feedback of what IM tactic might work best, candidates are most likely motivated to focus on their qualities and experiences for the job when writing the cover letter and CV and thus use more self-promotion compared to other IM tactics. Additionally, other tactics are also likely used less because of inappropriateness or incompatibility of tactics and format. Relatively strong norms exist concerning the content and format of a CV (Brown and Campion, 1994), providing less room or opportunity for the candidate to use a variety of different IM tactics and may foster the use of some IM tactics over others. Candidates may be able to implement self-promotion on their CV when listing or describing past (work) experiences. Ingratiation may be less possible to be incorporated on CV but more possible on the cover letter, as the latter format allows more freedom and contains less constriction on what content the cover letter is allowed to have. Past research shows that the less restricted a selection format is, the more diverse IM tactics can be (e.g. Stevens and Kristof, 1995). Therefore, the most freedom to use a variety of IM tactics contains the job interview, compared to CV and cover letter. The interview also gives the candidate the chance to potentially explain past

behavior that might not be perceived as favourably, such as gaps in the employment history. Such behaviors are referred to as defensive IM tactics.

Thus, we propose:

*Hypothesis 2a:* Candidates engage mostly in self-promotion and ingratiation in their CV and cover letter compared with other IM tactics.

*Hypothesis 2b:* Candidates use self-promotion, ingratiation, as well as defending one-self and making excuses during the job interview.

Past studies show that the assessment method can determine the variety of IM tactics shown but also the amount of IM tactics (McFarland et al., 2003). For instance, DeShon and Alexander (1996) found that when a person is concentrating on a task that is cognitively tax-ing, he or she is less able to self-monitor adequately. With less self-monitoring, the individual is less capable to use IM (accurately). Overall, past findings suggest that IM use may vary de-pending on how many cognitive resources an assessment method requires, i.e., the more cog-nitive resources the method requires, the less IM is implemented (McFarland et al., 2003). More precisely, if a method requires a great amount of an interviewee's cognitive resources, it is more difficult for that interviewee to devote attention to choosing IM tactics. In contrast, if the job candidate has a less demanding task during which he or she can navigate all his focus and attention to what is required, he or she has more cognitive resources available to choose IM behaviors.

A job interview, especially when structured with situational and biographical ques-tions, is a high demanding task for a candidate to navigate. The candidate does not only have to focus on the sometimes challenging and difficult questions, but also simultaneously think of appropriate answers and monitor adequate behaviors. In contrast, when he or she writes a cover letter and CV, the only task he or she has to accomplish is to make a good impression and be invited to an interview. Additionally, there are fewer time constrains in which he or

she has to finish this task and additional tools can be employed to accomplish the task successfully. We therefore argue that due to the candidate having more cognitive resources available when writing a cover letter and CV, the candidate will engage in more IM. More precisely, we propose:

*Hypothesis 3:* The total amount of IM tactics by a candidate will be higher in his/her cover letter and CV compared with his/her IM behavior during the interview.

**Competitiveness and IM behavior**

An additional variable that might influence IM behavior and in which both selection situations (cover letter/CV vs. selection interview) differ, is the perceived competitiveness to get the job and therewith the need to impress. When a candidate prepares his or her cover letter and CV, he or she does so mostly in isolation and with adequate time. The candidate might assume that several others apply for the job, but can only guess whether the other candidates are better suited for the job and how many apply. Thus the experienced competition is low and not salient to the candidate, it nevertheless is present in each selection step. Roulin, Krings, and Binggeli (2016) argued that the experienced competition could determine whether and to what extent a candidate fakes during personnel selection. Similarly, other researchers argue that the more competitive a person is, the more he or she does whatever it takes to impress (Tett et al., 2006). Past studies argued that although the experienced competitiveness lies within the person, it might be more or less activated by the situation. Roulin et al. (2016) argued that people differ in their competitive worldview (CW). People high on CW believe that the world is filled with competition where only the strongest survive and to be ruthless is at time necessary to win (Roulin & Krings, 2016). The lesser people score on CW the more they believe that the world is fair and competition is not needed to succeed in life.

Roulin and Krings (2016) see CW as an individual difference. During personnel selection, candidates high on CW fake more on personality questionnaires than candidates low on

CW (Roulin et al., 2015). Roulin and Krings (2016) argued that one's CW is stronger when the competition is more salient and it in turn affects faking behavior to a higher extent (Roulin & Krings, 2016). CW has so far only been addressed as an antecedent of faking behavior on personality measures (e.g., Roulin & Krings, 2016). The importance of CW within personnel selection is therefore present and seemingly influential but research extending CW to other personnel selection is rare (Roulin, Krings, & Binggeli, 2016). The current research aims to shed more light onto the role of CW within two selection steps: cover letter/CV and selection interview.

We argue that the influence of CW on IM behavior depends on the situation. Only guessing whether and how many other candidates are invited to an interview when writing a cover letter and CV means that competition is less salient and the influence of a candidate's CW is weaker compared with the face-to-face election situation, e.g., interview. Often candidates gather before a job interview in the same room and are invited into the interview room one after another. Thus, candidates see each other, interact with one another, and in turn know how much competition is present. We therefore argue that in such a situation, the competition is more salient and in turn, the influence of CW on IM is stronger. When such competition is obviously present, one's CW is more prevalent and candidates do whatever is necessary to win, including increased IM implementation (Roulin, Bangerter, & Levashina, 2015; Roulin & Krings, 2016; Roulin, Krings, & Binggeli, 2016). More precisely, we propose:

*Hypothesis 4:* The correlation between CW and IM behavior will be stronger during interviews compared with IM behavior in cover letters and CVs.

However, the situation and the salience of competition can not only influence a candidate's CW. Past research also shows that the situation may influence the expression of personality (e.g., Beaty, Cleveland, & Murphy, 2001; Chatman, 1989; Hattrup & Jackson, 1996). In turn, personality can determine how much a candidates uses IM during personnel selection.

**Big Five Personality Factors and IM Throughout the Selection Process**

A highly discussed moderator of the relationship personality and performance is the situation in which the task is taken place (e.g., Barrick & Mount, 1996; Beaty, Cleveland, & Murphy, 2001; Chatman, 1989). Mischel (1977) suggested that the situation could be placed on a continuum between "weak" and "strong". A weak situation exists of very few or no cues concerning how to behave or what is expected. Thus, the individual is without guidance as to what is appropriate or inappropriate. On the other hand, a strong situation provides specific clues on what is expected and what behavior is likely successful and appropriate. Mischel (1977) extended this knowledge and suggested that a strong situation constrains the expression of personality. If a person has a clear understanding of how to behave appropriately in a given situation, his or her personality does not play a major role. However, the person's personality has a larger impact during weak situations. Accordingly, individual differences can determine the outcome of the situation, whether it leads to success or failure.

A few studies have investigated the relationship between personality and situational strength in the work context. For example, Beaty et al. (2001) found that depending on situational strength, the relationship and contextual performance differed. Similarly, Meyer, Dalal, and Bonaccio (2009) suggested that personality determines performance outcome more strongly in weak situations. We argue that participating in a job interview corresponds to a weak situation compared with writing a cover letter and CV, which would be considered a strong situation.

When a candidate writes a cover letter, he or she often has specific guidelines of what a cover letter and CV should include (e.g., work experience, education, and motivation for the job). Therefore, the situation is stronger compared with a job interview. During the job interview, the candidate not only has to go through a more demanding and cognitively more draining task, but also has to decide within a few minutes how to appropriately react/answer the

questions. The candidate has to not only consider how he would react in certain situations (i.e., situational interview questions) or describe past behavior (biographical interview questions), but also decide what behavior is appropriate and how to describe it. The candidate has no time to carefully prepare his or her answers. When writing a cover letter he or she has plenty of time to choose what to express and how. Thus, a job interview provides a much weaker situation; the requirements are less clear and the appropriate behavior less visible. We argue that the candidate's personality influences IM more severe in a weak situation compared with a strong situation.

More precisely, we propose:

*Hypothesis 5:* Personality will predict IM behavior during a job interview compared but will not predict IM on cover letter or CV.

## Methods

### Participants and Procedure

Participants were informed through E-Mail, flyers and posters that they can enroll for an assessment center training. After enrollment each participant was informed three weeks prior to the assessment center training that they can send in a cover letter and CV for a management trainee position. A specific job advertisement was attached to that information e-mail. When e-mails were sent out to inform participants about the assessment center training, they were also asked to fill out two questionnaires (Big Five Personality Test, competitive worldview)

Seventy-eight participants (65.4% female, *M* age = 26.50, *SD* = 4.00) sent in their cover letter and CV two weeks before going through a one-day assessment center training.

Most participants were native German speakers (90.5%) and already completed a Bachelor Degree (47.4%).

Before sending in their application material, participants were informed that they would receive detailed feedback by an HR manager specialized in personnel selection. Within the assessment center training, the first exercise participants went through was a structured interview. Afterwards, participants had an individual session in which the HR professional provided feedback on their cover letter and CV. At the end of the day, participants received detailed feedback on their performance during the interview. During the assessment center training, participants were in contact with each other. The contact to other assessment center training participants served to increase the competitive and realistic assessment atmosphere.

**IM tactics in Cover letter, CV, and Interview**

Our approach to code IM tactics was based on earlier studies (e.g. Barrick & Mount, 1996; Peeters & Lievens, 2006). Two work and organizational psychology students coded all cover letter, CV, and interview IM behaviors. Both students completed three hours of training to successfully identify IM tactics. Specifically, they were provided with clear definitions and examples of each IM tactic. Before reading actual cover letters and CVs, they went through several mock cover letters and CVs to gain common understanding of the different IM tactics. Afterwards, discrepancies were discussed to ensure that category definitions were understood equally well among coders. Each coder coded all interviews and application materials independently. Inter-rater reliability reached $\alpha = .91$.

Coders summed how often a person expressed an IM tactic (either written in their CV and cover letter or verbally during the interview). The composite of assertive IM tactics consisted of the sum of the frequencies of self-promotion (description with superlative word, e.g. exceptional, outstanding, passionate, and adjective use, e.g. efficient, experienced, confident),

exemplification (giving specific examples of past work examples, e.g. " My most recent work project won an award"), and ingratiation (describing the hiring company in a flattering way, e.g. "The company's reputation stands for itself", or expressing ones own desires in reference to the company, e.g. "I would be thrilled to work for such a company"). The composite of defensive IM tactics included the frequencies of justification, and apologies.

## Interview and Materials

**Structured Interview.** Six professional trained interviewers conducted the interview. However, only one interview was active, i.e. asked the questions. The remaining interviewers were passive, i.e. only listened to and scored answers. The interview lasted approximately 30 minutes and consisted of nine situational and nine biographical questions. An example for a situational question is "Imagine the following situation: You and a co-worker work on a project together with the deadline is coming up. Your co-worker is not holding up his end of the bargain and does not provide his work results on time as agreed. What would you do?" An example for a biographical question is "Two of your co-workers or friends were in a fight and individually tell you about it and ask for your advice. In your opinion, one person was actually in the right. What was the situation, how did you behave, and what was the outcome". While participants answered the questions, interviewers took notes. All interviews were recorded on video for the purpose of IM coding afterwards.

**Competitive Worldview.** We used the 20-item Competitive Social Worldview scale ($\alpha$ = .79; Duckitt et al., 2002). Example items are "It's a dog-eat-dog world where you have to be ruthless at times" or "Winning is not the first thing, it is the only thing". Participants indicate their agreement or disagreement on a 7-point Likert scale, with 1 = strongly disagree and 7 = strongly agree.

**Big Five Personality Test.** We used the HEXACO-60 scale to assess participants' personality traits (Ashton & Lee, 2009). For the purpose of the study, we focused on only five

of the six dimensions that are included in the HEXACO (i.e. Openness to experience, Conscientiousness, Extraversion, Agreeableness, Emotionality/ Neuroticism, we excluded the honesty dimension). Participants used a computer to answer all 60 items on a 100-point Likert scale, with 0 = strongly disagree and 100 = strongly agree. Participants were able to position their agreement or disagreement anywhere on the scale and the computer registered the exact point (e.g., 50.2). An example item indicating openness to experience is “I would enjoy creating a work of art, such as a novel, a song, or a painting”. An example item indicating conscientiousness is “I pan ahead and organize things, to avoid scrambling at the last minute”. An example item indicating agreeableness is “I rarely hold a grudge, even against people who have badly wronged me”. An example item indicating extraversion is “I prefer jobs that involve active social interactions to those that involve working alone”. An example item indicating emotionality/neuroticism is “I sometimes can’t help worrying about little things”. All personality scales reached good reliabilities (openness to experience $\alpha = .66$; conscientiousness $\alpha = .74$; agreeableness $\alpha = .66$; extraversion $\alpha = .84$; emotionality/neuroticism $\alpha = .78$).

## Results

Table 1.1 provides an overview off all means, standard deviations, and correlations between measures.

The first goal of the current study was to shed more light onto the IM behavior of candidates during their initial contact with a potential new employer. Hypothesis 1 states that candidates already show IM tactics in their cover letters and CVs. Indeed, on average a candidate showed 38.35 IM behavioral incidence in their cover letter and CV, of which 37.90 were assertive IM behaviors and only a very small portion were defensive. This supports our first

hypothesis and underlines that IM behavior already plays a significant role when candidates apply for a job.

Looking more closely at what IM tactics were used, we can conclude that self-promotion was indeed the most prominent IM tactic used in cover letters and CVs (cover letter *m* = 12.49; CV *m* = 20.67; vs. interview *m* = 1.69). During the interview, candidates showed more variety in their IM behavior. During the interview, candidates showed self-promotion (*m* = 3.34) and exemplification (*m* = 4.96 ) the most, followed by justification (*m* = 3.16) and to a small extent intimidation (*m* = 1.71), ingratiation (*m* = 0.35), and apologies (*m* = 0.72). However, it is important to note that those tactics were used between one and two times on average throughout the interview. Overall, we can conclude that candidates do indeed engage in several different IM tactics during the interview, while focusing mostly on self-promotion when writing a cover letter and CV, thus supporting hypotheses 2a and 2b.

Our third hypothesis states that candidates use a higher amount of IM tactics when writing a cover letter and CV compared with participating in a job interview. On average, candidates engaged in 23.16 IM tactics during the interview. A paired-sample t-test reveals that the difference between IM tactics in cover letters/CVs and IM tactics used during interviews is indeed significant (t(77) = -15.19, *p* = 001). The test reveals that candidates engaged in more IM tactics in their cover letter and CV compared with during an interview.

The correlation between competitive worldview and IM tactics in cover letters and CVs, and the correlation between competitive worldview and IM tactics during interviews did not reach significance. We therefore cannot support hypothesis 4.

Next, we entered all five personality traits as predictors and the total amount of IM tactics used in cover letters and CVs as the dependent variable when conducting a linear regression. Table 1.2 provides an overview of all regression weights. The regression revealed that none of the Big Five personality indicators reached significance ($F(5,72) = .50$, $p = .78$).

However, when all five personality indicators were entered as predictors for the total amount of IM tactics used during the interview, the picture changed ($F(5,72) = 3.78$, $p = .01$). Agreeableness ($\beta = -.37$, $t(77) = 3.34$, $p = .01$) and emotionality/neuroticism ($\beta = .31$, $t(77) = 2.95$, $p = .01$) reached significance. The higher candidates scored on agreeableness the less they engaged in IM behaviors during the interview. However, candidates who scored high on emotionality/neuroticism showed more IM during the interview. We can therefore conclude that during a rather weak selection situation, compared with a rather stronger selection situation, candidates' personality plays a more significant role.

## Discussion

IM during personnel selection has been studied widely and intensively (e.g., Barrick & Mount, 1996; Barrick, Shaffer, & DeGrassi, 2009; Ellis, West, Ryan, & DeShon, 2002a; Kleinmann & Klehe, 2010). From such studies, we can draw the conclusion that IM does have an effect on the impression a candidate makes. Some studies confirm that when IM is used appropriately and at the right time, it can help to make a favorable impression and increase the chances of getting the job (McFarland, Ryan, & Kriska, 2003; Peeters & Lievens, 2006; Stevens & Kristof, 1995). Other studies showed that depending on the individual and the timing of IM, it can also backfire and create a more negative impression and therewith lower the chances of getting hired (Kacmar, Delery, & Ferris, 1992; Rudman, 1998). Due to such effects studying IM in all selection contexts is relevant.

However, a candidate use of IM to increase his or her chances of being hired was mostly studied with the job interview or a few assessment center exercises. Before a candidate gets to those selection steps, another selection step precedes. The first task that most candi-

dates have to accomplish is writing a cover letter and CV. Only if they make a good impression, they are invited for an interview. The current research highlights that candidates do engage in IM in cover letters and CVs, especially with self-promotion. However, whether candidates used IM and what tactic candidates employed, does not determine whether the same candidate engages in IM during a job interview. Therefore, even though both situations have the same motivation in common, i.e., to make a favorable impression, the situations are different enough to cause IM to a higher or lesser extent and may determine what IM tactic is implemented. We showed that while candidates mostly engage in self-promotion in cover letters and CVs, they use more of a variety of tactics during the job interview. Additionally, candidates use more IM when applying for the job in written form than when participating in an interview. Although it appears that, the situation determines how much IM is used, the CW a candidate holds seems to play a lesser role.

Researchers have argued that even though each individual has a general disposition on how competitive the world is, this disposition can be more or less activated depending on the situation (e.g., Roulin & Krings, 2016). In a competitive situation, peoples' competitive worldview increases in general but even more so among individuals who already score high on CW, and less so among people who score low on CW.

Our study fails to support the argument that the situation can activate candidates' CW. Neither writing a cover letter and CV nor participating in the interview revealed a relationship between CW and IM. When a candidate writes a cover letter or CV, he or she undertakes an (educated) guess on how many other individuals are likely to apply for the same job. Due to uncertainty concerning how competitive the market or the attractiveness of the job to other candidates, the candidates imagined competition might actually be as high as the observable competition when going through an interview. In our study, candidates saw how many others went through an interview. Thus, it is possible that the assumed difference between situations

was cancelled out because of the competition that the candidate imagined existed when writing the cover letter. So far, research has been rare on the relationship between perceived competitiveness and the usage of IM. However, especially when applying for the job and preparing one's application, candidates often assume that the market is harder or more competitive than it actually is. Additionally, some individuals think they have a better chance wining the competition than others (Dickerson & Taylor, 2000). For example, women often think they have a lower chance, not only because of feared discrimination but also because of the lack of self-esteem (Kling, Hyde, Showers, & Buswell, 1999). On the other hand, a more objective variable that would influence the perceived condition when applying for the job (especially when writing a cover letter) is the actual employment market. For example, Germany has a strong economy at the moment and thus more jobs to offer than other economies, e.g., Greece, or Spain. It would be interesting whether individuals' CW in countries with a weak economy is in general higher or more activated than candidates in countries in which the job market is stronger. Thus, future research should not only examine how competition influences the use of IM, but also what mediator or moderators can play a role, such as self-esteem, economics, and current (un)employment rates.

Individual differences that this study examined in reference to IM were the big five personality traits. In line with past research, this study shows that personality predicts the usage of IM (Beaty et al., 2001), however, not in all selection situations. We found that agreeableness and emotionality only influences IM usage during an interview, but not when writing a cover letter or CV. People who are more agreeable engage in less IM and people who are more emotional/neurotic engage in more IM. As we know, IM can influence on how an individual is perceived and in turn can influence whether the job offer is presented. Nevertheless, if an organization seeks an agreeable employee but one who is also low on emotionality, the

candidate the organization is looking for might actually engage in less IM, and thus may potentially lower his or her chances of getting a job offer. Past research also highlighted that people who score high on emotionality often choose IM tactics that are less successful, or achieve a "too much of a good thing" effect (Fletcher, 1990). Future studies could test how candidates who are especially high on emotionality use IM and how it interferes with interview questions or exercises aimed at testing emotionality or how emotionally stable an individual is. Researchers could therewith test whether the lack of IM usage of agreeable candidates interacts with their performance on tasks that test agreeableness or competencies highly related to that personality trait.

Overall, we can conclude that IM behavior starts earlier in the application process than investigated thus far. Candidates do engage strongly in IM. For all jobs, IM behavior can already influence who is invited for an interview, increasing one's chances of getting the job. However, depending on the job, the lack of IM might actually be caused by the personality trait that is valued and desired for a particular job. Hence, practitioners and researchers have an interest in how IM takes place during the very first selection step, what individual differences lead to engaging in IM behaviors, how objective variables can influence how much IM is used, and what the outcomes might be.

**Study Limitations and Directions for Future Research**

This line of research was one of the first attempts to shed more light onto IM behaviors in cover letters and CVs. While it shows that candidates do indeed engage in IM during the first initial contact to the employer, the study also bears some limitations.

Even though all candidates were asked to direct their application materials to a specific job advertisement, we did not control whether candidates actually did or whether they send materials that were addressed to other jobs. It is possible that depending on the job advertise-

ment candidates are inclined to engage in more or different IM tactics. Similarly, the interview´s format can determine to what extent candidates engage in IM. Future studies can therefore explore whether the listed qualifications, or tasks in the job advertisement (or the phrasing thereof) influence IM behaviors expressed on cover letters and CV. Additionally researchers can include yet another tool used by organizations to screen candidates, i.e. online application tools/forums. Many organizations have an online application system, which the candidate uses to provide information about his or her past work-experience, qualifications and motivation. In other words, often organizations specifically ask candidates not to send in a cover letter and CV but to fill out all information online. Researchers can therefore focus on how such online application tools influence IM behaviors. As past research shows, a structured interview minimizes IM behavior compared to unstructured interview. It is therefore possible that different formats on how to provide biographical data can also influence the extent of IM behavior.

Our study aimed not only to study whether candidate use IM on cover letter and CV but also whether a relationship exists between IM behavior on cover letter and CV and job interviews. Our study could not show that candidates who engage in cover letter and CV also do so during the job interview. However, this lack of relationship might have been caused by the structure of the interview. Such a structure causes that the personality or tendencies of the candidate play a lesser role, or can be expressed to a lesser extent. Therefore, it is possible that candidates who use IM on cover letter and CV will also do so during an unstructured interview. In other words, the unstructured interview gives enough room for personality or disposition to unfold, compared to the unstructured interview, which controls such personality effects. Similarly, due have the (free) choice on what and how to describe yourself on a cover letter, personality or candidates' general IM inclination might influence IM behavior. Future research can therefore focus on the relationship between application format (cover letter/CV, online application sys-

tems, and (structured interview). Possible the less structured the formats the stronger the relationship, i.e. the more the IM behavior carries over to the next selection step, as it depends rather on the candidate that the situation.

Overall, not only the IM research domain benefits from knowledge derived from studies that focused on IM during cover letter and CV but also practitioners can base their decisions on how to gather biographical data from their candidates. In some jobs or in some industry the ability to implement IM behavior successfully (one is perceived as likable and competent) might even be desired. Then organizations might decide to have candidates send in their cover letter and CV. However if research shows that an online application tool, or specific wording and phrasing of a job advertisement can minimize IM behavior , organizations can chose to implement such tools (or phrase their job advertisement differently) if IM behavior is not desired.

## Appendix A: Summaries of Statistics of Study 1- Impression Management

Table 1.1

*Means, Standard Deviations, and Correlations of IM Tactics, Competitive Worldview and Big Five Personality Traits*

| IM cover letter/CV | *M (SD)* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. IM tactics total | 38.35 (12.05) | | | | | | | | | | | | | | | | | |
| 2. Self-Promotion | 33.17 (11.69) | .98** | | | | | | | | | | | | | | | | |
| 3. Ingratiation | 3.90 (2.31) | .23* | .05 | | | | | | | | | | | | | | | |
| 4. Exemplification | 0.83 (0.98) | .10 | .04 | -.08 | | | | | | | | | | | | | | |
| 5. Justification | 0.41 (0.65) | .04 | -.03 | .08 | -.03 | | | | | | | | | | | | | |
| 6. Apologies | 0.04 (0.19) | -.21 | -.22 | -.17 | .10 | .08 | | | | | | | | | | | | |
| IM interview | | | | | | | | | | | | | | | | | | |
| 7. IM tactics total | 23.16 (8.51) | -.09 | -.06 | -.13 | .03 | -.12 | -.03 | | | | | | | | | | | |
| 8. Self-Promotion | 3.38 (2.11) | .09 | .10 | .01 | .10 | -.27* | .01 | .28* | | | | | | | | | | |
| 9. Ingratiation | 0.35 (0.52) | -.03 | -.04 | .04 | .04 | -.13 | -.13 | .20 | .15 | | | | | | | | | |
| 10. Exemplification | 5.00 (2.14) | -.09 | -.07 | -.10 | -.06 | -.01 | .01- | .34** | .12 | .22 | | | | | | | | |
| 11. Justification | 3.16 (2.19) | -.01 | -.01 | -.80 | -.02 | .15 | -.09 | .40** | -.04 | -.17 | -.01 | | | | | | | |
| 12. Apologies | 0.72 (0.92) | -.02 | -.01 | -.04 | -.18 | -.07 | -.09 | .17 | .01 | -.06 | -.06 | .12 | | | | | | |
| Competitive Worldview & Big Five Personality Traits | | | | | | | | | | | | | | | | | | |
| 13. Competitive Worldview | 2.54 (0.57) | -.03 | -.05 | .03 | .02 | .18 | .24* | -.05 | -.08 | -.33** | -.10 | .18 | .10 | | | | | |
| 14. Openness to Experience | 56.60 (6.44) | .02 | .03 | .04 | -.24* | .03 | -.13 | .03 | -.14 | .04 | -.07 | -.12 | -.01 | -.12 | | | | |
| 15. Conscientiousness | 63.72 (5.93) | -.02 | -.01 | -.13 | .15 | -.01 | -.02 | .04 | -.09 | .03 | .04 | -.12 | .26* | -.19 | -.02 | | | |
| 16. Extraversion | 52.63 (7.40) | -.03 | -.03 | -.01 | -.02 | -.01 | -.02 | .04 | .16 | -.05 | .27* | -.19 | -.19 | -.08 | .19 | .22 | | |
| 17. Agreeableness | 48.33 (5.78) | .16 | .13 | .13 | .01 | -.07 | .08 | -.31** | -24* | .09 | -.21 | -.29** | -.08 | .01 | .22 | .17 | .02 | |
| 18. Emotionality/Neuroticism | 46.01 (7.37) | -.06 | -.04 | -.12 | .08 | -.08 | .06 | .28* | -.08 | .13 | .11 | .11 | -.07 | .17 | -.09 | -.01 | -.01 | .04 |

*Note*. ** p < .01, * p < .05.

Table 1.2

*Predictors of the Linear Regression Including all Personality Traits*

| | IM cover letter/CV | | | IM use interview | | |
|---|---|---|---|---|---|---|
| Predictor (*n*=78) | *B* | *SE* | β | *B* | *SE* | β |
| Openness to Experience | -.05 | .23 | -.03 | .18 | .15 | .14 |
| Conscientiousness | -.10 | .25 | -.05 | .16 | .16 | .11 |
| Extraversion | -.03 | .20 | -.02 | .01 | .13 | .01 |
| Agreeableness | .37 | .25 | .18 | -.54 | .16 | -.37* |
| Neuroticism | .11 | .19 | -.07 | .31 | .12 | .31* |

*Note*. IM cover letter/CV $R^2$= .03; IM Interview $R^2$= .21; * p < .05

# Chapter 2:

# Miming the Ideal Employee: The Occurrence and Structure of the Ideal-Employee Factor During Personnel Selection.

Vanessa Jacksch & Ute-Christine Klehe

Justus-Liebig-University Giessen

Filip Lievens & Filip de Fruyt

Ghent University

## Author's Note

Vanessa Jacksch & Ute-Christine Klehe, Department of Work & Organizational Psychology, Justus-Liebig-University

Filip Lievens Department of Personnel Management and Work and Organizational Psychology, Ghent University; Filip de Fruyt Department of Developmental, Personality and Social Psychology, Ghent Univeristy

## Abstract

Candidates often distort their responses on personality questionnaire items. Besides the well documented score inflations on desirable items, this distortion causes a second phenomenon in the form of an additional "ideal-employee factor" (IEF) underlying responses of items from diverse personality dimensions. Less explored is how the IEF evolves and what pattern it has. The first study tests the emergence of the IEF among true applicants, i.e., whether applying for different jobs leads to different personality items loading onto the IEF. The second experimental study shows that different job profiles can indeed predict how an applicant scores on the big five personality items and, therewith, how the IEF is formed. It further addresses and supports the relevance of the individual and situational antecedents of response distortion as proposed by McFarland and Ryan (2000).

*Keywords:* response distortion, ideal-employee factor, personnel selection, personality assessment

# Miming the Ideal Employee: The Occurrence and Structure of the Ideal-Employee Factor During Personnel Selection

> "If the world operates as one big market, every employee will compete with every person anywhere in the world who is capable of doing the same job. There are lots of them and many of them are hungry."
>
> Andy Grove

Finding a job is a competitive activity, rewarding applicants who come across as suitable, hardworking, and thus promising future employees. To test whether applicants possess the desired qualities, organizations often rely on personality inventories, particularly inventories of the Big Five personality dimensions, i.e., conscientiousness, emotional stability, openness to experience, extraversion, and agreeableness (Barrick, Mount, & Judge, 2001; Ones, Dilchert, Viswesvaran, & Judge, 2007; Tett & Chrisitanen, 2007). While most scholars agree on the conceptual usefulness of the Big Five under non-evaluative conditions, they agree far less on the usefulness of testing them during personnel selection (e.g., Morgeson et al., 2007a, 2007b). Vulnerability to applicants' response distortion during personnel selection has caused a rather heated debate among professionals. After all, unlike the "normal" conditions under which personality tests have been developed and have most often proven their proposed five-factor structure (e.g., Costa & McCrae, 1985; Hogan & Hogan, 1992; Smith, Hanges, & Dickson, 2001; but see also Bäckstrom, Björklund, & Larsson, 2009), selection situations seem to cause a different structure of all personality traits. Some researchers have argued that the structural change is mainly caused by applicants not being rewarded for presenting their "true self", but for presenting a "self" that will get them the desired job (Mueller-Hanson, Heggestad, & Thornton, 2006; Hogan, Barrett, & Hogan, 2007). Therefore, applicants tend to

answer personality items in a way they think is valued positively by the hiring organization (Klehe et al., 2012).

The consequences of such distortion are severe. First, applicants who distort their answers well may have a better chance of getting the job. Particularly, response distortion leads to an inflation of scores, especially on the dimensions of conscientiousness and emotional stability (Birkeland et al., 2006), as applicants aim to present themselves as hardworking, responsible, and stress resistant. Second, the construct validity of the personality inventory suffers. Thus, the Big Five personality items not only load onto their respective personality traits, but also onto a sixth factor, which Schmit and Ryan (1993) labeled the "ideal-employee factor" (IEF). Many scholars have replicated this finding, particularly under selection conditions or when directly asking participants to present themselves as an ideal applicant for a job (Topping & Gorman, 1997; McFarland & Ryan, 2000; Hogan et al., 2007). While Bäckstrom et al. (2009) demonstrated that response distortion can also occur under 'normal' conditions, they too argue that this distortion is due to social desirability, i.e., the desire to make an overall good impression.

Despite its relevance, we still know relatively little about applicant response distortion. For one, most studies conducted in this area asked participants to actively distort their responses. Only a few studies to date have investigated actual response distortion among real job applicants; however, those studies rarely had a control group with which the distortion could be compared (e.g., Ellingson, Sackett, & Conelly, 2007; Hogan et al., 2007). Further, we need to extend our knowledge concerning the relationship between score inflation on the one hand, and the emergence of an IEF on the other. Testing for this relationship will be relevant, as score-inflation, the classic measure of response distortion, is a very difficult measurement to get during personnel selection since assessors usually lack the 'honest' scores with which to compare applicants' self-presentation during selection. If, however, score inflations

and the IEF represent the same phenomenon, we then might be able to use the IEF as a tool to identify the applicants who most eagerly distort, just as well as the items that are explicitly susceptible to such distortion. Further, if score inflation and the IEF represent the same phenomenon, then the same individual and situational variables that have been proposed to enhance score inflation may in turn also predict the emergence of the IEF.

The goal of this line of research is to address these fundamental questions. In the first study, with the help of two different groups of actual job applicants, we first illustrate the evidence of the existence of the IEF. With the help of the person-situation-interaction perspective (Fiske & Taylor, 1991; Mischel & Shoda, 1995, 1998), we argue that the composition of the IEF changes depending on the job opening. In the second study, we first replicate the findings of Study 1, and second, further show that the IEF and score inflation indicate the same phenomenon: applicants' response distortion. Finally, we rely on the theory of planned behavior (Ajzen, 1985) and the model of faking (McFarland & Ryan 2000, 2006) to identify individual and situational differences that predict response distortion. With this, the two studies make the following three contributions. First, we test whether the IEF pattern changes depending on the job. We will show that applicants distort differently between jobs to match their answers on the personality questionnaire to the desired traits the job entails. Second, we will show that the emergence of an additional sixth factor is the same as score inflation on the Big Five Personality Test items under a selection condition compared with a normal testing condition. Third, we will demonstrate that response distortion is influenced not only because of the situation the test taker is in, but also by the test taker's individual disposition.

## The IEF and Changes in Patterns Depending on Job Profile

A long tradition of research supports the notion that personality can largely be divided into five factors (the Big Five), and that these personality traits can help predict performance on the job (e.g., Tett & Christiansen, 2007). Other scholars, however, question the usefulness of the Big Five, particularly when assessed during personnel selection (e.g., Morgenson et al., 2007).

Schmit and Ryan (1993) found that the proposed five-factor structure indeed evolved in normal, non-threatening situations. However, during personnel selection, the classic five-factor structure failed to reflect the actual pattern of participants' responses; while all items continued to load onto their respective personality factors, some also loaded onto a sixth factor. This sixth factor then showed applicants to be very conscientious, productive and hard workers, very likable, courteous, and self-reliant - in short, the ideal employee.

Given that several studies have replicated the emergence of a sixth factor or unaccounted scale-interrelations among applicant samples (e.g., Biderman, Nguyen, 2009; Biderman, Nguyen, Mullin, & Luna, 2008; Burns, Christiansen, 2007; Cellar, Miller, Doverspike, & Klawsy, 1996; Collins & Cleaves, 1998; Ellingson et al., 1999; Klehe et al., 2012; Pauls & Crost, 2005; Topping & O'Gorman, 1997; Van Iddekinge, Raymark, Eidson, & Putka, 2001), we propose:

*Hypothesis 1:* Under selection conditions, a six-factor solution, including an additional factor combining items from different personality dimensions, will provide a better fit to data than the five-factor solution proposed by test developers.

## Explaining the Structure of the IEF: Interaction Between Person and Situations

So far, we have treated the IEF as a generalized phenomenon spanning across items addressing all personality dimensions. Bäckstrom et al. (2009) argued that most Big Five

questionnaire items are socially desirable; yet, past research on score inflation has shown that not all items are inflated in the same manner. Rather, while score inflation has been shown to be rampant in classic laboratory studies asking participants to 'fake good', studies addressing score inflation among real job applicants generally show far weaker and more refined effects. Here, the score inflations for most Big Five dimensions were small if not negligible (Ones, Viswesvaran, & Reis, 1996). Still, substantial inflations emerged for the personality dimensions conscientiousness and emotional stability (Birkeland et al., 2006). Similarly, Klehe et al. (2012) found that these two dimensions loaded the highest onto the IEF.

The reason for these more refined findings may lie in a person-situation-interaction perspective (Fiske & Taylor, 1991; Mischel & Shoda, 1995, 1998). The idea is that situational cues (e.g., a selection situation) can activate different mental representations or schemata within the person. Mental representations or schemata are cognitive structures that include and integrate affective reactions (i.e., emotions) and inferred traits for behavior in certain situations, thus guiding the processing of incoming information and providing scripts on how to react appropriately in this situation (Fiske & Taylor, 1991).

The basic argument underlying the IEF currently is that different situations and test purposes can activate different schemata within the same respondent: Low-stakes settings (e.g., voluntary participation in an anonymous study) are more likely to activate a "stranger-description" schema (Schmit & Ryan, 1993), i.e., a self-referenced evaluation of the perceived fit between personality items and the dominant self-schema (Holden, Kroner, Fekken, & Popham, 1992). Thus, the respondent describes him/herself most accurately to somebody unknown, i.e., a stranger. In such situations, responses usually fit the proposed Big Five factors structure rather well (Digman, 1990; Goldberg, 1990).

Yet, more evaluative settings, such as when applying for a job, may activate something very different. During personnel selection, applicants may be less concerned with pre-

senting themselves as accurately as possible, but instead may be more concerned with making a favorable impression and getting the job (e.g., Hogan et al., 2007). Consequently, the schema activated may be one of an "ideal-employee" (Schmit & Ryan, 1993). This ideal-employee schema integrates knowledge and assumptions regarding the desired traits of a qualified applicant, and thus guides behavior (e.g., responses given to the questionnaire).

Of particular relevance in this context might be the social desirability associated with different personality dimensions. While all personality dimensions may be desirable to some degree, some personality dimensions might appear particularly relevant, irrespective of the job applied for (Bäckstrom et al., 2009). Applicants might think, for example, that employers will prefer applicants with a strong work ethic, i.e., conscientiousness, as well as applicants who handle stress well and stay calm in moments of uncertainty, i.e., who show emotional stability (Huffcutt, Conway, Roth & Stone, 2001). Therefore, applicants want to present themselves especially favorably on those two dimensions independent of which job they apply for at the time. In line with this argument, past research on response distortion indicates that score inflation is particularly high on conscientiousness and emotional stability compared with smaller inflations for the remaining dimensions. Thus, the following hypothesis is proposed:

*Hypothesis 2:* Measures of conscientiousness and of emotional stability will load higher on the IEF than measures of less obviously desirable Big Five personality traits.

Yet, while conscientiousness and emotional stability might be desirable across jobs, the person-situation-interaction perspective would suggest that applying for different jobs might also activate quite different schemata, which in turn will guide behavior and responses, leading to response distortion on different dimensions. Thus, applicants for a position that requires interaction and expressing helpful behavior might think of a rather different schema

than those applying for a position in which they work in isolation (Holland, 1997). An applicant for a bus driver position, for example, might not only want to look reliable (i.e., conscientiousness), resilient, and able to handle stress, but, when thinking about explaining transfers, helping tourists find their way and assisting elderly or disabled passengers, the applicant might also want to show his or her high level of agreeableness and extraversion. When applying for a position in which one works in isolation, agreeableness and extraversion may appear less relevant. Thus, we propose:

*Hypothesis 3a:* The structure of the IEF will change depending on the job in question.

*Hypothesis 3b:* When applying for a job that requires interaction with and helping others, the measures of extraversion and agreeableness (besides conscientiousness and emotional stability) will also load higher on the IEF than the remaining personality traits.

*Hypothesis 3c:* Loadings of extraversion and agreeableness onto the IEF will be higher in the applicant group that applies for a job that requires interaction with and helping others compared with the applicant group that applies for a job that does not require these skills.

## Methods

### Participants and Procedure

Data was obtained from a large Belgian transit organization. Prior to participating in a selection interview, 7,083 job applicants for either a bus driver or metro driver position completed the Big Five Personality Test (90.7% men; bus driver $N$ = 6,351; metro driver $N$ = 732). No applicant received feedback on their answers to the questionnaire, and there was no time limit for completing the personality items.

**Job profiles and the Big Five Personality Traits.** Applicants for both jobs found the job advertisement on several online job agent websites and in local newspapers. The job advertisement entailed a job description and specifications that were developed with the help of O*Net 'work styles' (O*Net, Employment and Training Administration, 2010) and Holland's vocational theory (1996).

*Bus drivers:* The job advertisement listed several required tasks, e.g., assist passengers, collect tickets, and report delays or accidents. The job specifications communicated that especially the traits of realistic, social, enterprising, and conventional were important for the job. Applicants read, for example, that communication with others and helping tourists and passengers in need were valued. Bus drivers were asked to be open to explaining routes and transfers, providing information to tourists, and assisting and helping people without having to be asked or prompted.

*Metro drivers:* The job advertisement listed several required tasks, e.g., operate vehicle, report delays or accidents, perform routine maintenance on equipment, and determine when and what kind of maintenance is needed. The job specification communicated that especially the traits of realistic, conventional, and enterprising were important for the job. Applicants read, for example, that hands on problem solving and following set procedures were desired. Metro drivers were asked to follow rules and understand the importance of them, and since they are partly responsible for the smooth ride and safety of their passengers, they were prompted to be creative in problem solving. Additionally, the job description mentioned that they would have to make announcements to passengers occasionally.

## Materials

**Big Five Personality Test.** The Big Five personality traits were assessed with the NEO-PI-R (Costa & McCrae, 1992). The NEO-PI-R consists of 240 items answered on a 5-

point Likert scale, reaching from “strongly disagree” to “strongly agree”. The NEO-PI-R assesses 30 facets, 6 for each personality trait. All personality scales reached high reliability (see Table 2.1).

## Results

Table 2.1 presents the reliabilities, means, standard deviations, and correlations of the Big Five personality traits within each group. All measures showed acceptable internal consistencies.

To compare different models and groups, each personality dimension is represented by four parcels. To compose parcels, we followed a semi-random approach, i.e., each parcel represents one item of each personality trait facet. Thus, each parcel represents all facets of the personality trait in question. For example, conscientiousness includes items assessing six facets of conscientiousness, i.e. self-efficacy, orderliness, dutifulness, achievement striving, self-discipline and cautiousness. Each conscientiousness parcel contains items of each of those six facets, although which exact items was allocated randomly.

Figure A1 illustrates the five-factor model and Figure A2 the six-factor model (including IEF) that we fitted to each group.

*Hypothesis 1:* To test whether a five-factor or six-factor model fit the data better, we fitted both models to each group with the help of confirmatory factor analyses in AMOS 22, compared the goodness of fit indices, and conducted model comparison tests. For that purpose, we used the $x^2/df$ ratio, which should ideally be under 3 (Byren, 1994, 1998), the incremental fit index (IFI), the comparative fit index (CFI), a Tucker- Lewis index (TLI), all of which should be at least .90 or higher, and the root square error of approximation (RMSEA), which should be at most .08 or lower. To test which competing model fit the data better, we

conducted an $x^2$-difference test ($\Delta x^2$). However, in recent years researches have recommended an investigation concerning the difference in CFI values ($\Delta$CFI; e.g., Cheung & Rensvold, 2002). Cheung and Rensvold (2002) suggested the cut-off point of .01, i.e., a difference in CFI values less than .01 indicates invariance. All model goodness of fit indices and comparison tests are listed in Table 2.2.

In the bus driver group, the five-factor model achieved an acceptable fit. However, the six-factor model including the IEF indicated an even better fit, as shown by the chi-square and CFI difference test ($\Delta x^2 < .001$; $\Delta$CFI = > .01). In the metro driver group, the outcome shows the same pattern. The five-factor model achieved a good fit, but the six-factor model was a significant improvement of fit ($\Delta x^2 < .001$; $\Delta$CFI = > .01).

Overall, results therefore support hypothesis 1, i.e., the six-factor model fits the data better than the five-factor model. Next, we examined the factor loadings onto the IEF within each group, and the difference between groups.

*Hypothesis 2 – Hypothesis 3:* First, we examined the parcel loadings within each group, i.e., examined whether parcels of conscientiousness and emotional stability loaded highest onto the IEF (Hypothesis 2). Second, we compared parcel loadings between groups, i.e., whether an additional personality trait indicator relevant for the job load onto the IEF (Hypotheses 3). All standardized parcel loadings are represented in Table 2.5.

Within each group, all parcels loaded significantly and highly onto four of their respective traits (see Table 2.2). The loadings of conscientiousness parcels failed to reach significance. However, in the bus driver group, most parcels loaded significantly onto the IEF. Two parcels of openness to experience did not load onto the IEF. Within the metro driver group, a similar picture emerged, as most parcels loaded significantly onto the IEF, i.e., two openness to experience parcels did not load onto the IEF.

Taking a closer look, within the bus driver group, the average (i.e., average score of all loadings per trait) parcel loadings of conscientiousness onto the IEF was λ = .757. This average in parcel loading was the highest among all parcel loadings onto the IEF. The second highest loading onto the IEF was emotional stability parcels (λ = .618), followed by extraversion parcels (λ = .378) and agreeableness parcels (λ = .372). According to the critical ratio test, loadings of emotional stability parcels were higher than loadings of extraversion parcels (z = -16.59, $p < 0.05$), agreeableness parcels (z = -21.32, $p < 0.05$) and openness to experience parcels (z = -35.90, $p < 0.05$), but lower than loadings of conscientiousness parcels (z = -3.97, $p < 0.05$). Loadings of conscientiousness parcels were higher than loadings of agreeableness parcels (z = 19.20, $p < 0.05$), extraversion parcels (z = 21.13, $p < 0.05$) and openness to experience parcels (z = 35.22, $p < 0.05$). Loadings of openness to experience parcels were lower than loadings of extraversion parcels (z = -20.36, $p < 0.05$) and agreeableness parcels (z = 15.51, $p < 0.05$).

Within the metro driver group, again factor loadings of conscientiousness onto the IEF reached the highest loading onto the IEF (λ = .767). The second highest loading onto the IEF were emotional stability parcels (λ = .595), followed by extraversion parcels (λ = .372) and agreeableness parcels (λ = .314). According to the critical ratio test, loadings of emotional stability parcels were higher than loadings of extraversion parcels (z = -4.99, $p < 0.05$), agreeableness parcels (z = -7.04, $p < 0.05$) and openness to experience parcels (z = -11.83, $p < 0.05$), but lower than loadings of conscientiousness parcels (z = -3.01, $p < 0.05$). Loadings of conscientiousness parcels were higher than loadings of agreeableness parcels (z = 9.26, $p < 0.05$), extraversion parcels (z = 7.84, $p < 0.05$) and openness to experience parcels (z = 14.11, $p < 0.05$). Loadings of openness to experience parcels were higher than loadings of agreeableness parcels (z = 5.12, $p < 0.05$) and extraversion parcels (z = -7.22, $p < 0.05$). Loadings of extraversion parcels and agreeableness parcels did not differ from one another (z = -1.66, $p >$

0.05). We can therefore support hypothesis 2, i.e., conscientiousness and emotional stability load highest onto the IEF compared to the remaining personality indicators.

To examine whether factor loadings onto the IEF actually differ between groups, we conducted a multi-group analysis in SPSS AMOS. Both the unconstrained model (i.e., parameters were allowed to be estimated freely and differently between groups), as well as the constrained measurement weight model (i.e., parameters were fixed to be equal across groups) received good fit indices (fit indices are provided in Table 2.2). However, the chi-square difference test showed that the unconstrained model fits the data significantly better than the constrained model ($\Delta x^2 < .001$), though the CFI difference test indicated that the unconstrained model does not fit the data better than the constrained model ($\Delta CFI = < .01$). Therefore, the following results should be interpreted with caution.

Factor loading comparisons revealed that groups do not differ between all factor loadings onto the IEF (emotional stability loading average $z = -2.16$ $p < 0.05$; conscientiousness $z = .94$ $p > 0.05$; agreeableness $z = -.76$ $p < 0.05$; extraversion $z = .19$ $p > 0.05$; openness to experience $z = -.26$ $p < 0.05$).

## Conclusion

Overall, the six-factor structure applies to applicants' personality indicators better than the five-factor structure. We therefore replicated past findings and showed that the IEF contributes to a better fit. Additionally, we illustrated that the pattern of the IEF is related to the specific job opening. We showed that both emotional stability and conscientiousness load significantly onto the IEF and achieved the highest loadings of all five personality indicators be-

tween both applicant groups. This supports our assumption that applicants often identify conscientiousness and emotional stability as specifically important personality traits and answer personality items accordingly.

Loadings of extraversion received the third highest loading onto the IEF in both groups, followed by agreeableness. However, the multi-group analysis did not reveal any difference in loadings of the Big Five personality traits between groups. It is possible that metro drivers found extraversion and agreeableness as valued traits, and ranked them as important as bus drivers did, as they would also be responsible for their passengers, and be required to make announcements. Metro drivers, even though they mostly work in isolation, might feel that they would have to help and assist their passengers in cases of emergency.

Overall, we can therefore conclude that the IEF is not only relevant within an applicant group, but that its pattern may change depending on the job opening. In the next study, we not only want to yet again shed light on the changes in the pattern of the IEF, but also detail the relationship between score inflation and the emergence of the IEF, and why some applicants seem to distort more successfully than others.

## Study 2

The first study focused on the emergence and pattern of the IEF in real applicants. The second study aims to replicate these findings and further examines how the IEF comes about with the help of two data points, i.e., under non-evaluative testing conditions and under selection conditions.

**Emergence of the Ideal-Employee Factor and Score Inflation**

Studies addressing response distortion usually ask participants to answer personality inventories at two points in time: once in a low-stakes or non-evaluative condition, and a second time under a high-stakes, real or imagined personnel selection condition. When asked under high-stakes conditions, an additional factor evolves. The first purpose of Study 2 is to test, how far the IEF identified in Study 1 is truly a result of the evaluative selection context. Therefore, we will again illustrate that a six-factor structure is accurate for applicants, but extend the findings of Study 1 by showing that a five-factor structure remains accurate under non-evaluative testing conditions. We propose:

*Hypothesis 4:* The factor structure of a Big-Five personality inventory will change depending on whether the inventory is administered under anonymous and non-evaluative conditions, or under 'selection conditions' when candidates have a particular job in mind for which they want to present themselves as favorably as possible.

*Hypothesis 4a:* Under non-evaluative conditions, a five-factor solution will provide a better fit to Big-Five personality data than a six-factor solution.

*Hypothesis 4b:* Under selection conditions, a six-factor solution, including a common ideal-employee factor, will provide a better fit to data than a five-factor solution.

Next to the emergence of an additional factor, one of the oldest and most troubling findings regarding personality tests during personnel selection is a significant score inflation when comparing the two assessments, where results typically show significant score inflations during personnel selection (Birkeland, Manson et al., 2006; Viswesvaran & Ones, 1999). However, researchers often only have either the 'honest' condition responses or the high-stakes condition responses. Additionally, in the high-stakes condition, participants are usually asked to imagine applying for *a* job, not a specific job in particular. Therefore, the current study will not only study the emergence of a sixth factor, but also its relationship to

score inflation by studying participants' responses under non-evaluative conditions, and while imagining they are applying for a specific job.

This idea has not been sufficiently tested, likely due to the different research traditions and designs employed in studies addressing score-inflation and the IEF (Biderman & Nguyen, 2004). Only Klehe et al. (2012) correlated participants' factor loadings on the IEF with an estimate of their score inflation. While their results were promising, Klehe et al. (2012) had no individual scores derived from a non-evaluative situation upon which to base their score inflations. Instead, they had to rely on the somewhat outdated norm values of the personality test. Thus, Klehe et al. (2012) were working with a proxy instead of an accurate assessment of each participant's score inflation. Finding a strong relationship between an accurate measure of respondents' score inflation and their factor score on the IEF, however, would not only confirm that both findings represent the same phenomenon in different disguises, but would also offer an alternative (and easily obtainable) avenue for assessing respondents' levels of socially desirable responding during personnel selection. Thus, we propose:

*Hypothesis 5a*: The Big Five dimensions will be inflated (the average mean of each dimension will be higher) under selection conditions compared with non-evaluative conditions.

*Hypothesis 5b*: There will be a strong positive relationship between participants' ideal-employee factor score and the degree of score-inflation emerging under selection conditions.

**Changes in IEF Patterns Depending on Jobs**

As in the first study, we propose that the IEF pattern will again change depending on the job a candidate applies for, but we also test the role of score inflation. In the second study, we again propose that conscientiousness and emotional stability will be the most relevant traits perceived by candidates independent of the job opening. We therefore propose:

*Hypothesis 6:* Measures of conscientiousness and emotional stability will (a) load higher onto the ideal-employee factor and (b) inflate the most compared with the remaining Big Five personality traits.

However, in the second study, we chose to test the difference in IEF patterns between jobs that mostly differed in candidates' need to be social, helpful, and friendly on the one hand, and creative, flexible, and open to new ideas on the other. When a person, for example, applies for a position in which listening to others, expressing empathy, and helping others is required, we expect agreeableness to play a more significant role than conscientiousness and emotional stability. In contrast, when a person applies for a job in which creativity and an open mind are desired, openness to experience may play a larger role than conscientiousness and emotional stability.

A candidate for a counseling position in a clinical setting, for example, might not only want to look reliable (i.e., conscientiousness), resilient, and able to handle stress (i.e., emotional stability), but, when thinking about helping people in need, might also envision themselves as being friendly, approachable, and trustworthy (i.e., agreeableness). When applying for a more creative or investigative job, however, agreeableness may appear less relevant than, for example, curiosity and an openness to learn, create, and experience new things. Thus, we propose:

*Hypothesis 7a:* Both score inflation and the structure of the IEF will change depending on the job in question.

*Hypothesis 7b:* When applying for a counseling-related job, measures of conscientiousness, emotional stability, and agreeableness will (a) load higher on the ideal employee and (b) inflate more compared with the other two personality dimensions.

*Hypothesis 7c:* When applying for an investigative position, measures of conscientiousness, emotional stability, and openness to experience will (a) load higher on the ideal employee and (b) inflate more compared with the other two personality dimensions.

**Predicting the Occurrence of IEF: Individual and Situational Predictors**

Response distortion can help candidates get the job when done well (Hogan et al., 2007; Mueller-Hanson, Heggestad, & Thornton, 2006), yet not all candidates distort their responses equally or equally as well. McFarland and Ryan (2000; 2006) attempted to explain this by proposing a model of response distortion, and argue that to distort successfully, three ingredients are crucial: situation, motivation, and ability. The theory of planned behaviors (Ajzen, 1985; 2002) argues that the situational antecedent, perceived difficulty, can predict whether people engage in an activity (Ajzen, 2002). In the context of responding to personality inventory items, the situation needs to make it appealing to distort, or more precisely, the individual has to have the impression that the situation will allow distortion without it being too difficult to do so successfully. Thus, when candidates believe that distorting their answers on the Big Five questionnaire successfully is rather easy, they are more likely to engage in that activity. Thus, we propose:

*Hypothesis 8:* The perceived difficulty of responding to the Big Five personality questionnaire will positively predict candidates' level of response distortion as indicated by (a) their level of scale inflation and (b) the IEF.

However, just as the situation has to provide the opportunity to distort answers, the candidates' motivation also plays an essential role. McFarland and Ryan (2006) argued that the domain or test outcome needs to be important to the response distorter, i.e., he or she needs to be motivated to do well. If such a motivation is lacking, no distortion behavior will be implemented (Ajzen, 1985, 2002). Therefore, we propose:

*Hypothesis 9:* Candidates' motivation to do well on the inventory will predict their level of response distortion as indicated by (a) their level of scale inflation and (b) the IEF.

Yet, even with the strongest intentions, if the individual is incapable and lacks ability, the behavior will fail to be successful. Behavior is usually the function of two things: one's motivation to show this behavior and his/her ability (Locke, Mento, & Katcher, 1978). Additionally, McFarland and Ryan (2000) argued that successful response distortion would require a certain ability to distort. Such ability usually implies two things: a cognitive component in the form of an understanding of what is being measured, and a behavioral component in terms of acting upon such understanding. Given that the latter is usually not very challenging on classic personality-tests (i.e., respondents need to mark one of several response-options), McFarland and Ryan (2000, 2006) proposed that the knowledge of what is actually measured will be a key indicator of candidates' ability to distort their responses. People who are better at identifying what is expected of them in a certain situation are also more likely to implement appropriate behaviors (Kleinmann, 1993). We would expect similar effects for candidates' implicit understanding of what is measured on their level of response distortion, and thus propose:

*Hypothesis 10:* Candidates' implicit understanding of what is measured will predict their level of response distortion as indicated by (a) their level of scale inflation and (b) the IEF.

To summarize, this study not only aims to examine the factor structure of the Big Five personality traits under normal and testing conditions, but also to illustrate how the structure changes according to the job opening. Further, we will explore the relationship between response distortion and the emergence of a sixth factor. To end the study, we also dig a little

deeper into the situation, motivation, and ability requirements that successful response distortion needs. In total, we hope to shed more light on the specific patterns and individual differences that lead to the formation of the IEF.

## Methods

### Sample and Procedure

The study was conducted in the context of the 2010 "testweek" at the Universiteit van Amsterdam. Participants were first-year psychology students, most of them Dutch (92%). The Universiteit van Amsterdam requires all first-year psychology students to participate in several studies within their first undergraduate year. Over several days spread over a period of two weeks, students completed several questionnaires, tests, and assessments. Within these testing sessions, participants of the current study ($N$ = 320; 75% female; average age = 21 years, $SD$ = 4,7) completed a Big Five questionnaire at two points in time and under two different conditions. First, participants responded under normal, anonymous "testweek" conditions. Second, a week later, participants completed the same questionnaire, but this time after receiving one of two different sets of instructions. Before filling out the Big Five questionnaire for the second time, participants were instructed to answer the subsequent personality questionnaire in a way that increases the likelihood of receiving an invitation for a selection interview. At the end of the study, participants were asked to reflect upon their reactions to the assessment situation, which served to assess both the manipulation check and the proposed predictors motivation, knowledge about what is measured, and perceived difficulty.

## Materials and Measures

**Candidate condition and job description.** Two different experimental conditions were created to simulate a selection situation. Depending on the experimental condition, a participant read one of two job descriptions. Both instructions reminded participants of the financial and career relevance of obtaining a good job after their studies, and asked them to imagine applying for one of two positions: a therapeutic assistant or a journalist assistant. Here, participants read a job description and specialization of the respective job, with information being based on the O*Net 'work styles' (O*Net, Employment and Training Administration, 2010) and Holland's vocational theory (1996). For example, participants who read instructions for a therapeutic assistant job were informed about required working styles, such as high stress tolerance (emotional stability), dependability (conscientiousness), and having a social orientation, i.e., preference to work with others rather than alone (agreeableness). Participants in the journalist assistant condition read, for example, that the job requires attention to detail (conscientiousness), high self-control (emotional stability), as well as being innovative (openness to experience). Additionally, five subject matter experts (SEM) read the job description. We asked the SEM panel to read and name personality traits that they thought were most relevant when applying for the job. They named the same personality traits as proposed by Holland (1996) and the online O*Net.

**Big Five Personality questionnaire.** The Big Five personality traits were assessed with the Vijf Persoonlijkheidsfactorentest (5PFT; Elshout, 1999). The scale of each factor consists of 14 items, answered on a 7-point Likert scale, ranging from "totally disagree" to "totally agree". An example of the openness to experience item is "I am developed, read a lot, and find intellectual and cultural aspects important" ($\alpha = .80$ under normal testing condition and .87 under selection conditions). An example of the conscientiousness item is "I find it important to conform to accepted norms." An example item of agreeableness is "I have empathy

for others, recognize the importance of other's difficulty, and consider their interests." An example item of extraversion is "I like to have people around me." An example item of emotional stability is "I am self-controlled. I do not lose control during emotional situations."

**Situation: Test difficulty.** The experienced test difficulty was assessed with two items developed by Arvey et al. (1990). An example-item is "This test was too easy for me" (reversed coded). Participants answered on a 7-point Likert scale, ranging from "totally disagree" to "totally agree."

**Motivation: Test motivation.** Test motivation was assessed with items developed by Arvey, Strickland, Drauden, and Martin (1990). Item examples are "I wanted to achieve the top-score on this test" and "I wanted to do well on this test." Participants answered on a 7-point Likert scale, ranging from "totally disagree" to "totally agree."

**Ability: Implicit understanding of what is measured.** A good measure for this may be candidates' ability to identify criteria (ATIC; Kleinmann, 1993). People who score high on ATIC are better at identifying what is expected of them in a certain situation. Once they act accordingly, ATIC then enhances their success during personnel selection (Kleinmann et al., 2011; Koenig, Melchers, Kleinmann, Richter, & Klehe, 2006). The assessment of ATIC was adapted from Kleinmann et al. (2006) to the personality domain. Participants read 18 items describing a certain trait and indicated on a 7-point Likert scale how relevant they thought the trait would be when working in the actual job. Item examples are "Being open for new ideas: Intellectual curiosity, need stimulation of new ideas" and "Warmth: Friendly, enjoys talking, likes to have interaction with many people on a personal level." Depending on the condition, different traits were relevant. Whether a participant scored high on ATIC was therefore assessed by examining the scores of items assessing emotional stability, conscientiousness (for all conditions), and then agreeableness when being in the therapeutic assistant condition, and openness to experience when being in the journalist assistant condition.

**Candidate response distortion.** Candidate response distortion was assessed via the IEF and score-inflation. The IEF was estimated via SPSS AMOS as a latent variable influencing all Big-Five measurements. Score inflation was assessed by regressing the scores of the selection conditions on the scores obtained under non-evaluative conditions, as well as by conducting several t-tests to compare the means assessed under non-evaluative conditions with those assessed under selection conditions.

**Manipulations check.** To test whether the manipulation of the two experimental conditions was successful, participants stated which job they had just 'applied' for: (1) journalist assistant, (2) therapeutic assistant, (3) others. In the journalist assistant job condition, 29 participants did not indicate the respected job. These participants were excluded from any analyses, leading to a sample size of 147 in the journalist assistant job condition. In the therapeutic assistant job condition, all participants indicated that they 'applied' for a therapeutic assistant position, resulting in 173 participants in the therapeutic assistant condition.

## Results

### Preliminary Analyses

Table 2.2 presents the reliabilities, means, standard deviations, and correlations of the Big Five personality traits under non-evaluative and selection conditions, as well as predictors proposed by McFarland and Ryan (2006). All measures showed acceptable internal consistencies. Consistent with past research, the responses on the Big Five personality dimensions correlated meaningfully under non-evaluative conditions (Barrick & Mount, 1991) and very highly under selection conditions (Schmit & Ryan, 1993). The correlations between the Big

Five dimensions assessed under non-evaluative conditions with the Big Five personality dimensions assessed under selection conditions are weaker, which is not surprising if response distortion really played a role and thus influenced the pattern of the Big Five dimensions.

*Hypothesis 4*: We proposed that the factor-structure of the Big Five assessment would change depending on the testing situation: under non-evaluative conditions, the test should reveal the classic five-factor structure (Hypothesis 4a), whereas under selection conditions, an additional common factor should emerge whose loadings span across parcels from different dimensions (Hypothesis 4b). Therefore, we compared two models in each condition (see Figures A1 and A2 in Appendix A). Model 1 assumes the Big Five personality traits, i.e., extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience, which were allowed to correlate with each other. To prevent analyses having to handle too many sets of variables (70 items loadings on five or six factors), we again computed three parcels per personality trait by calculating the averages of four or five items each, while trying to distribute items with high vs. low factor loadings onto the respective personality dimension as equally as possible across parcels (Hall, Snell, & Foust, 1999). Model 2 mirrored the first model, with only an extra latent construct added, which was allowed to affect all 15 item-parcels and was uncorrelated to any of the Big Five personality dimensions.

As expected, the five-factor model fit the data under normal testing conditions acceptably in each condition (for fit-indices see Table 2.4). The six-factor model fit the data under normal testing conditions also acceptably well, however non of the loadings onto the sixth factor reached significance. Under the selection conditions, the five-factor model did not fit the data well, particularly in terms of a suboptimal $x^2/df$ ratio and RMSEA value. As expected, however, adding a latent IEF significantly increased the fit of the model (therapeutic condition: $\Delta x^2 = 111.40$ (14), $p < .01$; $\Delta$CFI, $p > .01$; journalistic condition: $\Delta x^2 = 66.40$ (14),

$p < .01$; $\Delta$CFI, $p > .01$). Parameter estimates loaded onto both their respective Big Five personality dimension as well as on the IEF (see Table 2.6), thus supporting hypothesis 4b and replicating findings of earlier studies (Klehe et al., 2002; McFarland & Ryan, 2000; Schmit & Ryan, 1996).

*Hypothesis 5a – 5b:* The next hypotheses assumed that the IEF emerging under selection conditions is essentially the same as score inflation under selection conditions. In other words, we expect that the IEF under selection condition correlates positively with the score inflations on the Big Five items. When testing for such score inflations via paired-sample t-tests, results revealed significant effects for the average Big Five score inflated under selection conditions compared with non-evaluative conditions ($t$ (319) = -24.306, $p < .001$), as well as for each of the separate Big Five dimensions. Participants scored higher on extraversion under selection conditions than under non-evaluative conditions ($t$ (319) = -8.52, $p < .001$), higher on agreeableness under selection conditions than under non-evaluative conditions ($t$ (319) = -15.505, $p < .001$), higher on conscientiousness under selection conditions than under non-evaluative conditions ($t$ (319) = -21.655, $p < .001$), higher on emotional stability under selection conditions than under non-evaluative conditions ($t$ (319) = -18.663, $p < .001$), and finally, scored higher on openness to experience under selection conditions than under non-evaluative conditions ($t$ (319) = -22.060, $p < .001$), thus supporting hypothesis 5a.

To show that the IEF emerging under selection conditions represents the same phenomenon as the inflation of scores, we expect that the IEF is able to explain additional variance in the Big Five traits assessed under selection conditions after controlling for their effect assessed under selection conditions.

With the help of AMOS, we imputed IEF factor scores for every participant. Correlations between this IEF score with all Big Five measures assessed under normal as well as selection conditions indeed showed that only some correlations between the IEF factor score

and Big Five assessments under non-evaluative conditions reached significance (e.g., extraversion, agreeableness, see Table 2.2). Under selection conditions, however, the correlations between the IEF factor score and all Big Five dimensions reached significance and were considerably stronger (see Table 2.2).

With the imputed IEF score, we conducted a multiple regression analysis. We computed the average score of all 70 Big Five personality items assessed under selection conditions, and another average score of all personality items assessed under non-evaluative conditions. We first entered the average Big Five items assessed under non-evaluative conditions as a control variable to estimate true score inflation. As a second predictor, we entered the imputed IEF score. Both predictors were regressed onto the average Big Five personality indicators assessed under selection conditions. The results of the regression indicated that the two predictors explained 85.3% of the common variance ($R^2 = .85$, $F(2) = 918.70$, $p < .01$). The Big Five personality indicators did not predict the Big Five personality indicators assessed under selection conditions ($ß = -.03$, $p > .05$), but the IEF was found to be a significant predictor ($ß = .93$, $p < .001$). These results provide support for hypotheses 5b, i.e., there is a strong relationship between the IEF and score inflation under selection conditions.

*Hypothesis 6 -7:* We assumed that under selection conditions, conscientiousness and emotional stability indicators would load higher on the IEF than the other three remaining Big Five personality indicators, whereas hypothesis 7b and 7c made similar assumptions about agreeableness and openness to experience indicators, depending on the job in question.

In the therapeutic counselor condition, all personality indicators loaded onto their respective latent constructs. However, all personality indicators also loaded onto the IEF. Average loadings of all three parcels per dimension on the IEF indicated that agreeableness indicators reached an average factor loading of .78 onto the IEF, emotional stability indicators a

loading of .74, and conscientiousness a loading of .56. Extraversion and openness to experience indicators also loaded onto the IEF, but reaching rather low factor loading (openness to experience = .48; extraversion = .28). According to the critical ratio test, loadings of emotional stability parcels were higher than loadings of extraversion parcels ($z = 2.87$, $p < 0.05$), and agreeableness parcel loadings were higher than extraversion parcels ($z = 2.59$, $p < 0.05$). The remaining comparisons between parcel loadings did not reach significance. Next to agreeableness, conscientiousness and emotional stability achieved the highest loadings, however only agreeableness and emotional stability reached a significant difference between other parcel loadings. Thus, at this point we can only support hypothesis 7a, 7b and hypothesis 7c partially for the therapeutic assistant condition.

In the journalist assistant condition, all personality indicators loaded onto their respective latent constructs as well as onto the IEF. Openness to experience reached the highest factor loading onto the IEF (.81), followed by emotional stability (.78) and conscientiousness (.76). In this instance, extraversion and agreeableness also showed high loadings onto the IEF (extraversion = .72; agreeableness = .73). According to the critical ratio test, only loadings of agreeableness parcels were lower than loadings of openness to experience parcels ($z = 2.43$, $p < 0.05$). At this point we can only partially support hypotheses 7b and 7c.

*Multi-group analysis with both groups:* To test whether factor loadings onto the IEF actually differ between groups, we conducted a multi-group analysis in SPSS AMOS. Both the unconstrained model, as well as the constrained measurement weight model received good fit indices (see Table 2.4). However, the chi-square difference test shows that the unconstrained model fits the data significantly better than the constrained model ($\Delta x^2 < .001$). Again, the CFI difference test indicates that the unconstrained model does not fit the data better than the constrained model ($\Delta CFI = < .01$).

Parameter comparisons reveal that within the therapeutic condition, at least two out of three agreeableness parcels loaded significantly higher onto the IEF compared with the journalist condition ($z$ = - 2.00 $p < 0.05$). Not all remaining parcels loaded significantly differently onto the IEF between groups.

*Testing faking antecedents—based on McFarland and Ryan's model (Hypotheses 8-10):* Hypotheses 8 to 10 addressed the relevance of different proposed predictors of candidate response distortion. To test the predictive value of each antecedent, we first conducted a regression analysis onto the participants' score under the selection condition, controlling for their score during the non-evaluative condition. Next, we ran a comparable model in AMOS to predict the IEF (Appendix, Figure A3).

*Regression analysis on score inflation:* We first entered the computed average of the Big Five personality indicators assessed under non-evaluative conditions as the independent variable. Secondly, we entered all predictors proposed by McFarland and Ryan (2000) as independent variables. All these factors were regressed onto the average Big Five indicators under selection conditions, thus the average inflated Big Five indicators. The regression as such yielded significance ($F(4) = 60.44$, $p < .001$). The Big Five personality loadings assessed under non-evaluative conditions reached significance ($t(4) = 3.35$, $p = .001$), as well as indicators of test taking motivation ($t(4) = 3.41$; $p = .001$), ATIC ($t(4) = 10.07$, $p = .001$), and perceived test difficulty ($t(4) = -3.20$, $p = .002$). These results give further support for the notion that response distortion causes score inflation, as well as supports hypotheses 8, 9 and 10. As a next step, we fitted a model, including all predictors, hoping it would reveal similar prediction patterns and reach an acceptable fit.

*SEM.* We fitted a model including all Big Five latent variables with their respective indicators and the sixth factor-IEF, plus all variables assumed to predict the emergence of the IEF (i.e., test taking motivation, ATIC, and test difficulty). Including all variables, the model

achieved a good fit (depiction of model see Appendix B3; for fit indices see Table 2.4). Looking at the predictors, ATIC (.67), motivation (.45), as well as perceived test difficulty (-.31) significantly predicted the IEF.

We can conclude that ATIC, test taking motivation, and perceived difficulty of the test are stable predictors of the IEF, score inflation, and thus response distortion.

## Discussion

Job applicants' response distortion on personality inventories is a concern for both research and the practical field (Allport, 1937; Griffith & Peterson, 2006; Viswesvaran & Ones, 1999). Practitioners are worried that this distortion will lead to offering a job to the wrong applicant, and that the estimation of an applicant's successful performance is inaccurate. Academics debate what the antecedents, mechanisms, and consequences of response distortion during personnel selection are (McFarland and Ryan, 2000; Morgeson et al., 2007b; Ones et al., 2007; Smith, Hanges, & Dickson, 2001).

The current studies tackle several points concerning response distortion and the emergence of the ideal-employee factor. First, it shows that under real job conditions, an additional factor emerges, and that the pattern of the IEF is not constant. It shows that emotional stability and conscientiousness always load the highest, and explores whether a third personality trait indicator might load onto the IEF that might be particularly relevant for a certain job. Second, the second study supports and extends these results by investigating whether the IEF is the same as score inflation of the Big Five personality indicators assessed under selection conditions. Third, the study informs us about certain individual differences and situational factors that foster distortion, explaining why some participants distort.

In total, this research demonstrates that the emergence of a sixth factor under selection conditions is essentially a common appearance that might be difficult to avoid. Both studies thus indeed replicate the findings of earlier studies (Biderman & Nguyen, 2009; Biderman, Nguyen, Mullin, & Luna, 2008; Burns & Christiansen, 2007; Cellar, Miller, Doverspike, & Klawsy, 1996; Collins & Cleaves, 1998; Ellingson et al., 1999; Klehe et al., 2012; Pauls & Crost, 2005; Schmit & Ryan, 1993; Topping & O'Gorman, 1997; Van Iddekinge, Raymark, Eidson, & Putka, 2001). Additionally, this line of research not only informs us of the existence of the IEF, but also its composition.

Overall, the current line of research underlines that under evaluative situations, such as personnel selection, the five-factor structure does not hold true, but that a six-factor structure provides a more accurate picture. Under selection situations, that additional factor can be rightfully labeled as the IEF, as the current line of research supports and extends past findings by showing that the IEF may most often consist of personality indicators that are most relevant for the work in general, i.e., conscientiousness and emotional stability, and an additional one that is relevant for that particular job. Our studies illustrate that the pattern of IEF is more predictable than researchers and practitioners previously assumed. Applicants who apply for a position with a social orientation; nurses may distort differently than applicants who apply for a position with a different focus, e.g. lawyers. In turn, different IEFs will form in each job area. By knowing how applicants may respond on a personality questionnaire, responses can be controlled for, or the effects thereof can be managed and may even be used as a selection tool.

When applicants are capable of correctly identifying what requirements or personality traits are needed in a specific job, they may also be able to correctly identify what is required to perform successfully. If an applicant's IEF is in accordance with the job's requirement, or desired personality traits, that applicant might be the right person for the job. However, the

current line of research not only shows that the formed IEF can be used to identify response distortion, but that score inflation is essentially the same as the formation of the IEF. The second study shows clearly that score inflation under selection situations is the same as the emergence of the IEF. Nevertheless, because past research has mostly focused on whether and how an additional factor evolves, no study to date has tested whether some applicants distort more or differently than others. By demonstrating high correlations between the IEF and score inflations under selection conditions, conducting a regression analysis onto the inflated scores with the IEF as a predictor, as well as demonstrating the similarity between the emergence of a sixth factor and similar score inflations throughout all hypotheses, we can conclude that the emergence of the IEF and score inflation of the Big Five personality indicators under applicant conditions are the same phenomenon. This extension of knowledge is an important contribution to the field of personnel selection. Response distortion during personnel selection is a common albeit worrisome occurrence. Studies investigating response distortion can now use either the emergence of a sixth factor and its factor loadings or score inflation under selection conditions as indicators of response distortion, as this study demonstrates that both are appropriate.

The second study extends our knowledge regarding predictors of the IEF. It demonstrates that applicants with the ability to identify criteria and the motivation to do well will distort more than applicants who lack such ability and motivation. Additionally, the second study highlights that the situation can also influence whether an applicant will distort responses (or to a lesser extent). If the person has the impression that the test is rather easy, he or she is more likely to distort responses. However, such perceived difficulty can be manipulated. Future studies could test whether the knowledge of later going through an exercise in

which their personality traits are observed would lead to less distortion. In our study, the difficulty to distort was perceived as low, which is one limitation that researchers could compensate for in future studies.

**Study Limitations and Directions for Future Research**

This line of research has its limitations. In the first study, it is possible that the job profiles were too similar in nature. Essentially, the only clear difference between bus and metro drivers is that one group works in isolation and the other group works among people, but also with minimal actual contact with people on some days. Thus, future studies may compare real applicant groups that are completely different from one another. For example, research could focus on the difference between metro drivers and service personnel. The former group does not need to be extraverted, while the latter group needs to be very extraverted. By collecting greater data on different job applicants who apply for essentially different jobs, researches could eventually be able to cluster job groups according to their applicants' scores or patterns of the IEF.

The setting of the second study allowed for high experimental control, e.g., presenting different job advertisements with comparable differences in required skills and abilities. At the same time, however, the setting inherited some disadvantages. First, all participants were psychology students. Psychology students may be used to filling out questionnaires, and have a deeper or more educated understanding of testing and its purposes. Second, even though the simulation of a selection situation was successful, knowing that they were actually not applying for a job might have diminished the effects of some factors, such as the perceived test difficulty. Still, significant contributions to our understanding concerning response distortion were made that we might be able to generalize into a larger context.

Another concern that can be addressed in the future is the absence of a control group. A control group, in which participants are asked to apply for *a* job, would have been helpful to test the change in the IEF to a greater extent. When having this comparison possibility, researchers are not only able to compare applicants who apply for different jobs with each other, but also compare these conditions with the condition in which an applicant is asked to apply in a way that gets him or her at least one job offer of some kind.

Future studies that have a primary focus on the emergence of the IEF should also include individual differences that can cause response distortion, increase or lower distortion, or make someone a better distorter, i.e., distorting in a way that makes a positive impression. A contentious debate exists around whether response distortion is something good, bad, learned, or common (Klehe et al., in press; Ones et al., 2007, Ones, Viswesvaran, & Reiss, 1996; Rosse et al., 1998; Tett & Christinasen, 2007). This study shows that simply simulating a selection situation causes response distortion, and thus distorted answers are given. Mueller-Hanson et al. (2006) showed that response distortion influences hiring decisions. However, whether simply response distortion increases the chance of being hired, or whether only appropriate response distortion increases such a chance, should be examined in the near future. This study shows that response distortion happens, an IEF emerges, and due to some individual and situational factors, some distort more, some less, and some do it better than others. Nonetheless, the question of what constitutes "better" remains. It is possible to argue that an applicant who accurately identifies the requirements of the job, and therefore distorts his or her answers accordingly on the Big Five questionnaire, is capable of not only correctly identifying requirements while applying for the job, but also distorting behavior in a way that leads to success while working in the actual job. Therefore, it could be debated whether an applicant's (correct) IEF, i.e., if pattern of the IEF matches job requirements, should actually be used as a selection tool.

Thus, when personality inventories are used to make hiring decisions, response distortion should be kept in mind. Still, more studies are needed to examine factors that influence distortion and how one actually distorts responses. Only when identifying factors that diminish the occurrence of faking, can we control it if necessary to do so, or use them to identify the right applicant.

## Appendix B.1: Summaries of Statistics, Model Indices & Parameters of Study 2-IEF:

Table 2.1

*Means, Standard Deviations, and Correlations of the Big Five Personality Traits IEF Study 1*

| Bus Driver ($N$ = 6351 ) | α | M (SD) | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|
| 1. Extraversion | .79 | 2.54 (.30) | | | | |
| 2. Agreeableness | .79 | 2.78 (.29) | .23* | | | |
| 3. Conscientiousness | .85 | 2.91 (.31) | .43* | .37* | | |
| 4. Emotional Stability | .86 | 1.39 (.35) | -.33* | -.36* | -.56* | |
| 5. Openness to Experience | .76 | 2.27 (.29) | .42* | .10* | .14* | -.05* |
| Metro Driver ($N$ = 732) | | | | | | |
| 1. Extraversion | .83 | 2.51 (.33) | | | | |
| 2. Agreeableness | .82 | 2.77 (.31) | .29* | | | |
| 3. Conscientiousness | .88 | 2.94 (.33) | .40* | .35* | | |
| 4. Emotional Stability | .89 | 1.32 (.39) | -.34* | -.36* | -.62* | |
| 5. Openness to Experience | .80 | 2.31 (.32) | .40* | .17* | .09* | -.10* |

*Note*. * $p < .05$

Table 2.2

*Means, Standard Deviations, and Correlations of the Big Five Personality Traits and McFarland and Ryan's Antecedents of IEF Study 2*

| Non-evaluative condition (N = 320) | α | M (SD) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Extraversion | .85 | 4.58 (.85) | 1 | | | | | | | | | | | | |
| 2. Agreeableness | .80 | 5.10 (.66) | .34* | | | | | | | | | | | | |
| 3. Conscientiousness | .78 | 4.54 (.73) | -.07 | .12* | | | | | | | | | | | |
| 4. Emotional Stability | .88 | 4.69 (.92) | .19* | .33* | .03 | | | | | | | | | | |
| 5. Openness to Experience | .74 | 4.39 (.64) | .20* | -.03 | .15* | .08 | | | | | | | | | |
| Selection Condition (N = 320) | | | | | | | | | | | | | | | |
| 1. Extraversion | .81 | 4.98 (.71) | .44* | .28* | -.01 | .16* | .13* | | | | | | | | |
| 2. Agreeableness | .87 | 5.82 (.70) | .08 | .27* | .07 | .08 | .06 | .33* | | | | | | | |
| 3. Conscientiousness | .86 | 5.53 (.76) | -.05 | .01 | .40* | .01 | .06 | .35* | .56* | | | | | | |
| 4. Emotional Stability | .88 | 5.78 (.76) | .05 | .05 | .09 | .24* | .10 | .31* | .68* | .58* | | | | | |
| 5. Openness to Experience | .86 | 5.36 (.73) | .05 | -.01 | .06 | -.01 | .35* | .47* | .47* | .64* | .59* | | | | |
| McFarland & Ryan | | | | | | | | | | | | | | | |
| 6. Test motivation | .81 | 4.94 (1.10) | .13* | .09 | .05 | -02 | ..02 | .21* | .30* | .36* | . 31* | .36* | | | |
| 7. ATIC | .77 | 5.40 (.63) | .10 | .15* | .18* | -.03 | .07 | .33* | .52* | .56* | .43* | .49* | .42* | | |
| 8. Perceived test difficulty | .86 | 3.32 (1.14) | -.13* | -.08 | -.06 | -.11 | -.16* | -.23* | -.22* | -.22* | -.26* | -.26* | -.16* | -.23* | |
| 9. IEF | | 3.66 (.41) | .30* | .23* | .10 | .13* | .16* | .84* | .67* | .64* | .73* | .36* | .55* | -55* | -.28 |

*Note*. * p < .05

Table 2.3

*Overview of Model Fit Indices of IEF Study 1*

| | Study 1 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Bus Drivers Condition | $x^2$(df) | $x^2$/df | NFI | IFI | TLI | CFI | RMSEA | $x^2$diff (dfdiff) | CFIdiff |
| Five-factor Model | 2825.29 (160) | 17.66 | .95 | .95 | .94 | .95 | .05 | 1109.47(19)* | .02** |
| Six-Factor Model | 1715.82 (141) | 2.17 | .97 | .97 | .96 | .97 | .04 | | |
| Metro Drivers Condition | | | | | | | | | |
| Five-Factor Model | 490.82 (160) | 3.07 | .94 | .96 | .95 | .96 | .05 | 123.96(19)* | .01** |
| Six-Factor Model | 366.895 (141) | 2.60 | .95 | .97 | .96 | .97 | .05 | | |
| Multi-Group Analysis | | | | | | | | | |
| Unconstrained Model | 1849.51 (307) | 6.02 | .97 | .97 | .97 | .97 | .03 | 81.13 (33)* | .00 |
| Constraint Model | 1930.64 (340) | 5.68 | .97 | .97 | .97 | .97 | .03 | | |

*Note*. Selected fit indices of each model fitted to data with SPSS AMOS 24.

* $p < .001$

** $p \geq .01$

Table 2.4

*Overview of Model Fit Indices of IEF Study 2*

| | Study 2 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Non-evaluative Therapeutic Condition | $x^2$(df) | $x^2$/df | NFI | IFI | TLI | CFI | RMSEA | $x^2$diff | (dfdiff) | CFIdiff |
| Five-Factor Model | 160.29(82) | 1.95 | .88 | .94 | .92 | .94 | .07 | 35.85(14)* | | .02** |
| Six-Factor Model | 124.44(68) | 1.83 | .91 | .96 | .93 | .96 | .07 | | | |
| Non-evaluative Journalistic Condition | | | | | | | | | | |
| Five-Factor Model | 157.21(82) | 1.92 | .86 | .93 | .91 | .92 | .08 | 47.05(14)* | | .04** |
| Six-Factor Model | 110.16(68) | 1.62 | .90 | .96 | .94 | .96 | .07 | | | |
| Select Therapeutic Condition | | | | | | | | | | |
| Five-Factor Model | 217.68(82) | 2.66 | .88 | .92 | .90 | .92 | .10 | 111.40(14) | | .05** |
| Six-Factor Model | 106.28(68) | 1.56 | .94 | .98 | .97 | .98 | .06 | | | |
| Select Journalistic Condition | | | | | | | | | | |
| Five-Factor Model | 176.15(82) | 2.15 | .90 | .94 | .92 | .94 | .09 | 66.40(14)* | | 02** |
| Six-Factor Model | 109.75(68) | 1.61 | .94 | .97 | .96 | .97 | .07 | | | |
| McFarland Model | 190.40(113) | 1.69 | .90 | .96 | .94 | .96 | .06 | | | |
| Multi-Group Analysis | | | | | | | | | | |
| Unconstrained Model | 258.15 (138) | 1.87 | .93 | .96 | .94 | .96 | .05 | 31.18 (20)* | | .00 |
| Constrained Model | 289.33 (158) | 1.83 | .92 | .96 | .95 | .96 | .05 | | | |

*Note*. Selected fit indices of each model fitted to data with SPSS AMOS 24.

* $p < .001$

** $p \geq .01$

Table 2.5

*Overview of Factor Loadings of Six-Factor Model IEF Study 1*

| Study 1 | | |
|---|---|---|
| Bus Drivers | Respective Personality Trait | IEF |
| 1. Extraversion | .60* | .38* |
| 2. Agreeableness | .60* | .37* |
| 3. Conscientiousness | -.14 | .76* |
| 4. Emotional Stability | .47* | .62* |
| 5. Openness to Experience | .68* | .10 |
| Metro Drivers | | |
| 1. Extraversion | .63* | .37* |
| 2. Agreeableness | .66* | .31* |
| 3. Conscientiousness | -.13* | .78* |
| 4. Emotional Stability | .38* | .59* |
| 5. Openness to Experience | .71* | .07 |

*Note*. Factor loadings onto their respective personality trait and onto the IEF.

* at least two out of the four item parcels reached significance, $p < .05$.

Table 2.6

*Overview of Factor Loadings of Six-Factor Model IEF Study 2*

| Study 2 | | |
|---|---|---|
| Therapeutic Select Condition | Respective Personality Trait | IEF |
| 1. Extraversion | .70* | .28* |
| 2. Agreeableness | .40* | .78* |
| 3. Conscientiousness | .63* | .57* |
| 4. Emotional Stability | .40* | .74* |
| 5. Openness to Experience | .61* | .48* |
| 6. ATIC | | .71* |
| 7. Motivation to do well | | .42* |
| 8. Perceived Test Difficulty | | -.29* |
| Journalistic Select Condition | | |
| 1. Extraversion | .35* | .73* |
| 2. Agreeableness | .32 | .73* |
| 3. Conscientiousness | .38* | .76* |
| 4. Emotional Stability | .35* | .78* |
| 5. Openness to Experience | .31* | .81* |
| 6. ATIC | | .61* |
| 7. Motivation to do well | | .49* |
| 8. Perceived Test Difficulty | | -.30* |

*Note*. Factor loadings onto their respective personality trait and onto the IEF.

* at least two out of the four item parcels reached significance, $p < .05$.

**Appendix B.2**

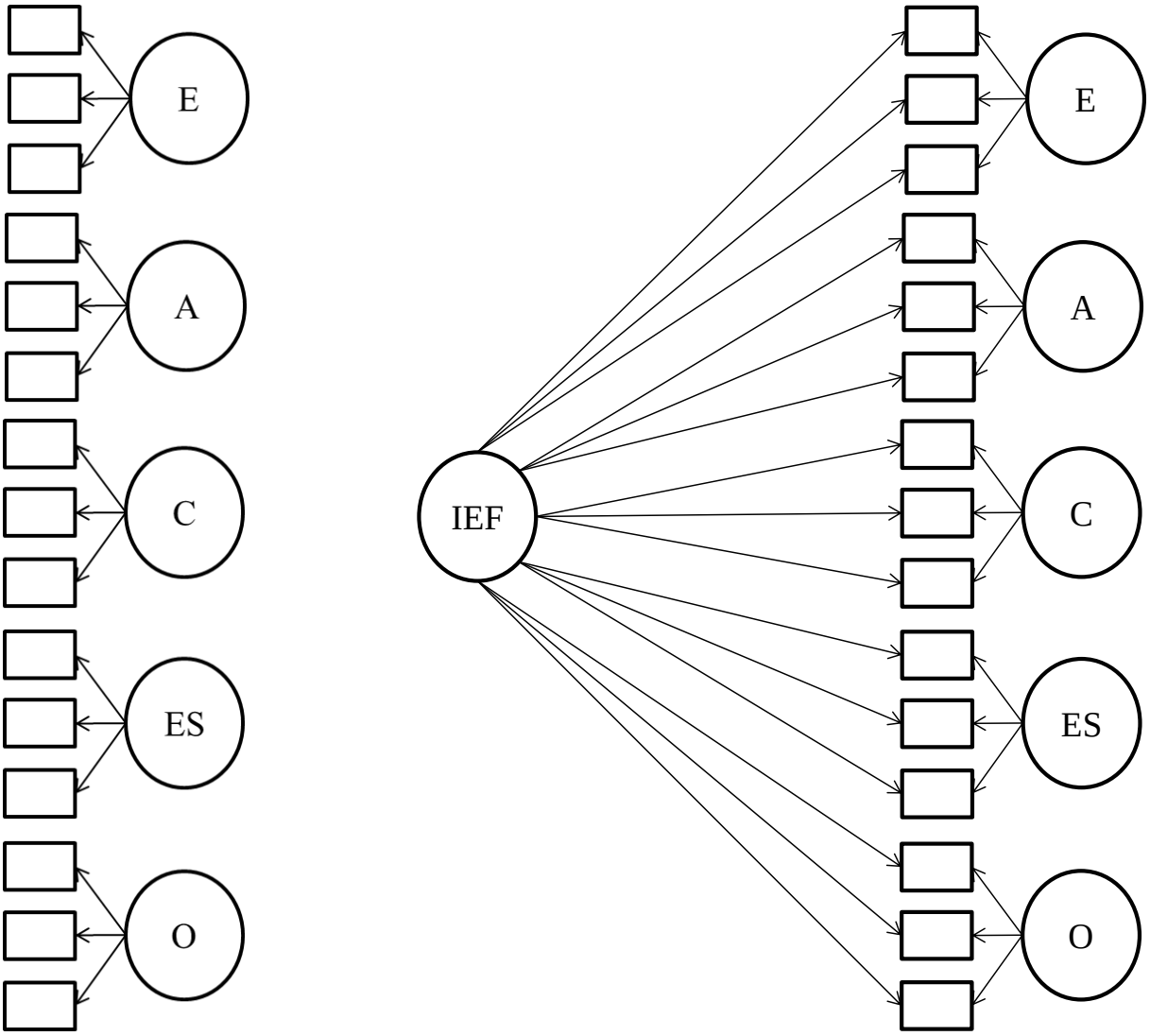


*Figure 1.A1/1.A2.* 5-Factor Model IEF Study fitted to Data

*Note*. Big Five Personality traits were allowed to correlate. In Study 1, each personality trait had four parcel indicators.

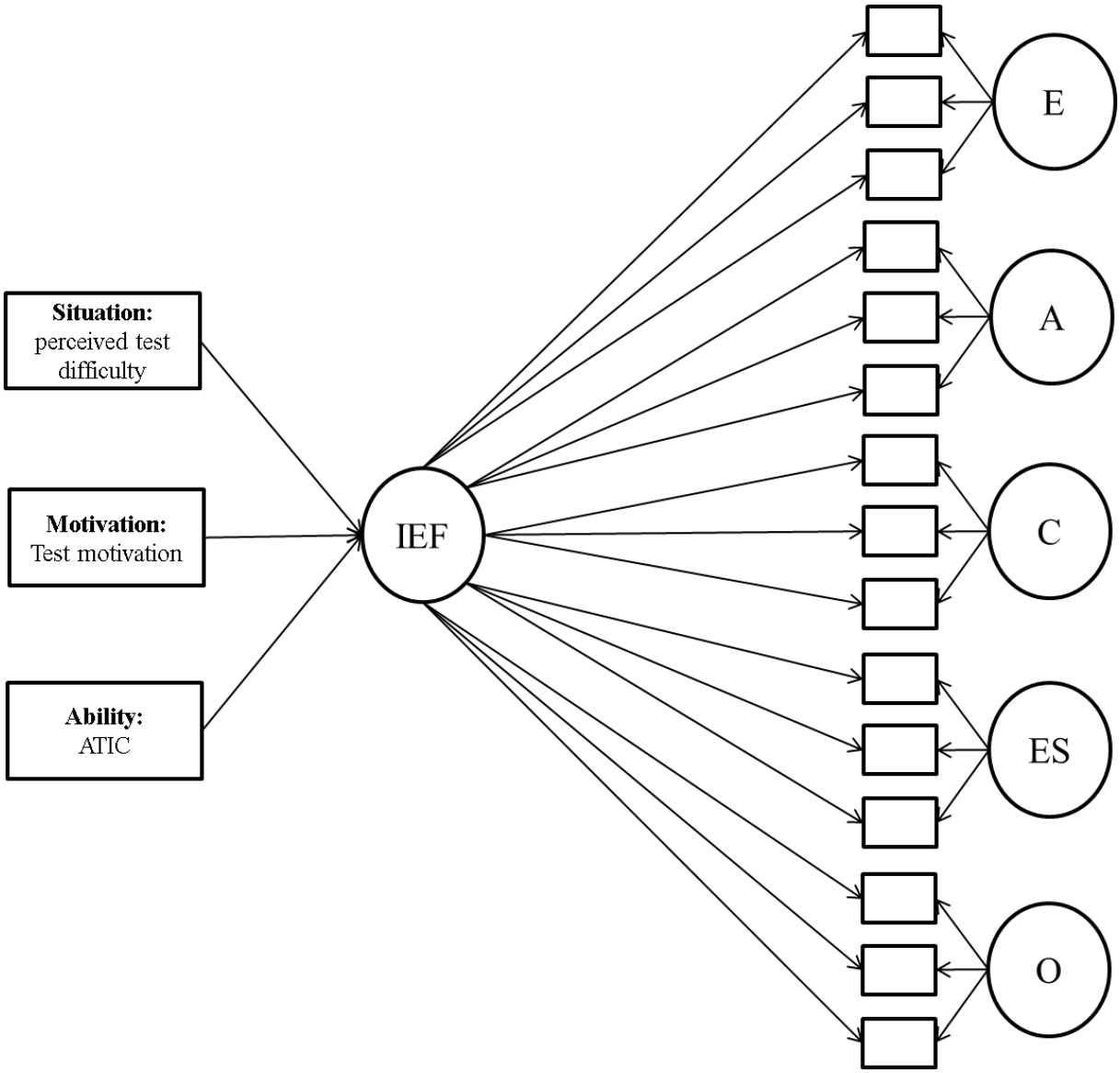


*Figure A3.* 6-Factor Model and Additional McFarland's Predictors fitted to Data

*Note*. Big Five Personality traits were allowed to correlate. In Study 1, each personality trait had four parcel indicators.

# Chapter 3:

# Unintended consequences of transparency during personnel selection: Benefitting some candidates, but harming others?

Vanessa Jacksch & Ute-Christine Klehe

University of Giessen

Author Note

Vanessa Jacksch, Department of Work & Organizational Psychology, Justus-Liebig-University; Ute-Christine Klehe, Department of Work & Organizational Psychology, Justus-Liebig-University

This research was funded by the German Research Foundation (DFG - KL 2366/3-1). We sincerely thank Kimberly Smith-Jentsch for providing the simulation employed in Study 1.

Published: Jacksch, V., & Klehe. U-C. Unintended consequences of transparency during personnel selection: Benefitting some candidates, but harming others? *International Journal of Selection and Assessment, 24,* 4-13.

## Abstract

Past research suggests that transparency during personnel selection procedures, i.e., revealing to candidates the dimensions on which their performance is being assessed, increases both fairness and candidates' performance. Two experiments challenge this assumption and propose that this effect only holds for non-threatening performance dimensions. Yet, when the dimension revealed targets candidates with a negative stereotype, their performance may suffer. In Study 1, both men and women performed better when they learned that a selection simulation targeted planning skills. Yet, when women learned that the simulation targeted leadership skills, they performed worse. Study 2 revealed a marginally significant interaction between transparency condition, gender, and stigma consciousness. In summary, transparency during personnel selection may thus benefit only some groups of candidates while hurting others.

*Keywords*: personnel selection, transparency, performance, stigma consciousness, stereotype threat

# Unintended consequences of transparency during personnel selection: Benefitting some candidates, but harming others?

Francis Bacon once said "Knowledge is power" (García, 2001). In other words: The more we know about the purpose, goals or expectations in a situation, the better we will succeed in dealing with it. Especially in uncertain, maybe threatening situations, knowledge offers situational control (Bargh, 1990) by providing actors with clues about what to do (Fishbach & Trope, 2005), thus minimizing the risk errors. Already at the age of three, humans appreciate situational knowledge to guide their behavior (Piaget, 1950) and people especially look for cues and control when they feel threatened with possible failure (Piaget, Brown, & Thampy 1985).

Personnel selection is such a situation in which failure is possible and in which people usually appreciate the situational control arising from knowledge. Consequently, scientists and practitioners in personnel selection often inform job candidates about a selection situation's requirements before sending them through the respective selection-task (e.g., Klehe, König, Richter, Kleinmann, & Melchers, 2008). Transparency arguably increases fairness among candidates and between groups. Knowing what is expected, i.e., understanding the process of decision-making, gives candidates a sense of procedural justice, i.e., the perception that the process of selection is fair and comparable across candidates (Folger, 1987; Gilliland, 1993) and therefore also enhances candidates' trust in the decision's accuracy (McFarlin & Sweeney, 1992). Indeed, when all candidates have the same information about the selection tool's criteria, then they all can focus on the relevant behavior, rather than some of them possibly misinterpreting the situation and concentrating on different and possibly wrong behavior (Kleinmann et al., 1996; Smith-Jentsch, 2007). As a consequence, transparency usually increases candidates' performance (e.g., Kolk, Born, & Flier, 2003) and improves the selection tool's construct related validity (e.g., Kleinmann, Kupsch, & Köller, 1996). According to this

rationale, selection tools should therefore always be made transparent. However, what if transparency may also cause exactly the opposite effect among at least some candidates, i.e., the loss of control over the situation, leading to self-doubt and failure?

**A potential downside to transparency**

"The greater the knowledge, the greater the doubt", Johann Wolfgang von Goethe once said (Douglas, 1917), expressing the opposite of what Francis Bacon said. Both statements might be true, though, and the current studies aim to identify circumstances that determine whether knowledge is power or causes doubt. The basic tenant of the following argument is that transparency does indeed facilitate performance under most circumstances, but that depending on the candidate and the information revealed, transparency may also cause stereotype threat. Stereotype threat is a quandary in which members of a certain social group are faced with the possibility of being judged or treated in a stereotypically consistent manner. Often then, these people become afraid of doing something that will be interpreted as stereotypically consistent with their group (Steele, Spencer, & Aronson, 2002). For example, Steele and Aronson (1995) told one half of a group of black American students that a test they were to take assessed their verbal ability. The other half of the group received no such information. There is a strong stereotype in the United States that Blacks are less intelligent than Whites, also with regard to verbal ability (Pfeffley & Hurwitz, 1998). Black students who received the prior information about the test content performed worse than white students or than black students who received no information (Steele & Aronson, 1995). A similar study extended this finding to women taking a math test. When women were reminded prior to test taking about the existing stereotype that women are worse in math than men, these women performed worse than women who were not reminded about this stereotype (Spencer, Steele, & Quinn, 1999). Interestingly, it did not matter in either study whether the black or female per-

son believed the stereotype to be true. The "threat in the air" to confirm to a stereotype existing in society is enough to cause individuals targeted by this stereotype to underperform (Smith, 2004). However, not only members of traditionally disadvantaged groups can be victims of stereotype threat, but stereotype threat can affect anyone, depending on the circumstances (Smith, 2004). For example, white participants being told about the better athletic skills of Blacks performed worse on a golf-playing task compared to a control group that did not receive such information (Stone, Lynch, Sjomeling, & Darley, 1999).

A stereotype consists of a target (i.e., social group) and a specific content, attribute, domain or any other factor that the stereotype judges (Steele & Aronson, 1995). In order for stereotype threat to occur, however, people further need to be in an evaluative situation where they are compared to others and they need to have the motivation to perform well (Schmader, Johns, & Forbes, 2008). Personnel selection usually satisfies these conditions: candidates are in an evaluative situation in which their performance is being compared to that of others, and they usually want to perform well. Yet, there is one more ingredient relevant for a stereotype threat to occur, namely the salience of the stereotype in regard to one's own group (Steele & Aronson, 1995; 1998). Transparency may cause this salience. The aim of the first study is to show that transparency in a single-actor selection simulation can indeed boost performance when the dimension revealed is stereotype free, but that the same transparency can harm the performance of candidates who are the target of some negative stereotype associated with this dimension. The second study aims to extend these results to a group-based simulation and additionally investigate the role of a common moderator of stereotype threat, i.e., stigma consciousness.

## Study 1

### Stereotype threat and its influence on performance

For stereotype threat to occur and to impair a person's performance, the person needs to be acutely aware of a stereotype that targets his or her social group (Stone & McWhinnie, 2008), i.e., the stereotype needs to be salient to the performer. Here, both subtle cues and blatant manipulations can activate stereotypes, i.e., make the stereotypes more salient, starting a circuit of doubt that eventually impairs performance (Smith, 2004). Blatant manipulations come in the form of telling participants prior to the test or exercise that their social group usually performs worse compared to other social groups on this type of test (e.g., that women tend to do worse on math tests than men or that Blacks tend to do worse on verbal ability tests than Whites; Steele & Aronson, 1995) or in explaining that the task at hand measures attributes or abilities about which stereotypes exist (e.g., that the following test assesses math abilities about which the society holds the stereotype that men are better in math than women; Stone & McWhinnie, 2008). Within personnel selection, such blatant manipulations are both unlikely and ethically unsustainable.

Subtle cues, however, may cause just the same (and according to Stone & McWhinnie, 2008 even stronger) stereotype threat and consequently impair performance just as much. Subtle cues include being the only female among male participants while working on potentially gender-sensitive material (e.g., math tasks; Inzlicht & Ben-Zeev, 2000), having a member of the comparison group administer the test (i.e., a male administrator when females were taking a math test; Marx & Goff, 2005), or informing participants about the content of a test (König & Eagly, 2005, Steele & Aronson, 1995). Informing participants about the test content or the underlying dimension of a selection procedure is exactly what transparency entails, and

thus, transparency might unintentionally cue stereotype threat and lower performance. Derived from stereotype threat theory, this would imply, for example, that the information that the next selection tool will assess leadership skills could activate the stereotype that females are worse leaders than males (Heilman, 2012). Female candidates would then experience stereotype threat. However, when men learn that leadership will be assessed, transparency will likely foster their performance.

Thus, we assume that the past findings, that transparency in personnel selection gives candidates more power to perform well, will generalize only to performance dimensions that are either neutral or are positively associated with one's social group. Here, transparency may well increase focus and the concentration on relevant information and may therewith improve performance (Kolk et al., 2003; Mischel & Shoda, 1995). Therefore we assume that:

*Hypothesis 1a:* When planning is made transparent, then both men's and women's performance will increase, compared to when planning is not made transparent.

*Hypothesis 1b:* When leadership is made transparent, men's performance increases compared to men's performance when leadership is not made transparent.

However, when transparency implies revealing information that targets a negative stereotype about the candidates' social group such as gender, performance may suffer. For women, transparency on the dimension leadership may thus not only clarify the requirements of the task but may also distract them by eliciting doubt and/or fear to potentially conform to a stereotype. Thus, we propose:

*Hypothesis 2:* When leadership is made transparent, women's performance suffers, compared to women's performance when leadership is not made transparent.

## Method

**Prestudy.** In order to verify that leadership is indeed connected to gender stereotypes and that planning does not contain any such gender stereotypes, we recruited 177 participants

(M age = 27.21, SD = 7.91, 107 female) via social network portals to fill in a short online questionnaire. Participants were asked to report on a 11-point Likert scale ranging from -5 to +5 (including the zero point) whether they think a gender stereotype exists within society about the performance dimensions planning and leadership and if so whether it favors men (negative numbers) or women (positive numbers). Irrespective of their own gender, participants rated the dimension planning neutrally, i.e., no gender stereotype exists concerning planning (*M* = .47, *SD* = 2.54). However, participants rated the dimension leadership as gender stereotyped, i.e., society stereotypes women to be worse leaders than men (*M* = -2.89, *SD* = 1.83).

**Participants and simulation.** One hundred and twenty-two graduate students and graduates (68 female, *M* age = 26.89, *SD* = 6.57) participated in a one-on-one personnel selection training. This training lasted between one and two hours in total and included both the selection simulation relevant for this study as well as subsequent feedback on both one's performance during the simulation as well as on one's resume and application letter. Also, participants could ask the female graduate HR-student running the experiment any information on selection procedures in general.

During the simulation participants were in the role of a customer representative in a hospital emergency room and interacted with a video-based scenario (Smith-Jentsch, 2007). Within the scenario, different characters appeared to speak directly to the participants and participants were instructed to react verbally to those characters. Each character provided different information, sometimes placing participants in a dilemma in which participants had to decide on the spot and defend decisions. During several trigger events, participants had to respond to unreasonable requests, demonstrate opinions, defend interests or structure sequences of activities. Scenarios ran continuously in that the computer automatically detected the participant's voice, recorded it and the character to whom the participant appeared to talk to was

looped, i.e., leaving the impression that the character was waiting for a response. If the participant did not say anything or stopped talking for a long enough time, the computer automatically continued with the next sequence.

During the recruitment, participants learned that this task is not a typical computer task but an interpersonal task more akin to classic assessment center tasks such as a series of role-plays. Thus, participants knew that no prior computer skills or knowledge were needed to perform well, as was indeed the case as the simulation needed no further input by the participants than their verbal responses to each scenario presented to switch to the subsequent scenarios.

**Manipulation and experimental design.** We employed a two (gender: female/male) by three (transparency conditions) between subject design in order to test our hypotheses. Participants were randomly assigned to one of three conditions: the non-transparent control condition (n=45), transparency on the gender-neutral performance-dimension planning (n=39), or transparency on the gender stereotyped performance-dimension leadership (n=48). Following the procedure employed by earlier interview and assessment center studies testing the effect of transparency (Klehe et al., 2008; Kleinmann et al., 1996), participants of all three conditions read the instruction for the simulation task before doing the actual simulation. In the non-transparent condition, participants received no information about the underlying dimensions. In the two transparent conditions, participants learned that the simulation served to assess their planning skills or their leadership skills, respectively.

**Performance.** Besides a degree in work and organizational psychology, the experimenter had gained experience in personnel selection prior to her involvement in the current study. While being blind to the study's research purpose and conditions, she had been explicitly trained for observing performance on the dimensions leadership and planning in the current simulation via a classic frame of reference training in order to ensure (1) a correct under-

standing of the dimensions leadership and planning, (2) attention to specific behaviors and understanding on how to observe them accurately and completely, and (3) a consistent standard for rating behaviors across candidates.

The observer employed a validated observation sheet (König et al., 2007) adjusted for the current study. This observation sheet contained examples of desirable and undesirable behaviors representative of both leadership and planning as well as blank space for notes on observations. The observer took extensive notes during the simulation and rated each dimension on a scale from 1 (undesired, unsuccessful behavior) to 5 (desired, successful behavior). Planning was operationalized as participants' ability to set priorities, structure information correctly and develop and gain a correct understanding of the situation. Leadership was operationalized as participants' assertiveness, emotional stability, stress resistance and decisiveness.

## Results

In order to test hypotheses 1a, we ran an independent factorial ANOVA on candidates' planning performance. There was no significant main effect of gender ($F(1,74) = 2.209$, $p = .16$ $\eta^2 = .028$) but a significant main effect of transparency condition ($F(1,74) = 42.03$, $p < .001$, $\eta^2 = .375$). Both men and women performed better on the dimension planning, when planning was revealed to them ($M = 4.0$, $SD = .62$) compared to when it remained non-transparent ($M = 2.71$, $SD = 1.03$). Therefore, this supports that when a dimension that neither targets men nor women with a (negative) stereotype is revealed, transparency can have a boosting effect on performance.

To test hypothesis 1b, we compared the mean of leadership performance between men who learned that leadership was assessed and men who did not receive such information.

There was no difference in leadership performance between men in the non-transparent condition ($M$ = 3.85, $SD$ = 1.04) compared to men in the transparent condition ($M$ = 3.79, $SD$ = .77). Therefore, the hypothesis could not be confirmed ($t$ (37) = -.21, $d$ = .29, $p$ = .84).

To test hypothesis 2, we compared the mean of leadership performance between women who learned that leadership was assessed and women who did not receive such information. In line with hypothesis 2, women performed worse on leadership in the transparent leadership-condition ($M$ = 2.6, $SD$ = .85) than in the non-transparent condition ($M$ = 3.20, $SD$ = 1.02, $t$ (52) = -2.13, $d$ = .59, $p$ =.04).

## Discussion

Together, results of Study 1 indeed support the notion that transparency can have different effects depending on the participants' gender and the dimension revealed. While transparency on the dimension planning proved beneficial to all candidates, transparency on the dimension leadership had a detrimental effect on women's performance. As no such inhibiting effect was found for men, findings from Study 1 challenge the notion that transparency is an essential component of a procedurally fair personnel selection (Gilliland, 1993). Rather, results suggest that transparency may inadvertently contribute to the low representation, of, for example, women in leadership positions. In order to shed more light on the potential negative effects of transparency more research is needed, especially with bigger sample sizes. In the current study the sample is rather small and as stereotype threat can be caused very easily but also diminished quickly, more studies are called for to find out when, how, and in what circumstances such a threat can occur under transparent selection conditions.

Additionally, the nature of the task (computerized simulation) may have also influenced participants' performance. Even though all participants have been informed beforehand that the simulation does not require any computer skills, it is still possible that some participants, and possibly women in particular, felt more threatened than men by having to work

with a computer. Additionally, this exercise was conducted individually, with only simulated interaction partners, and performance was rated by only one observer, albeit a well-trained one. The next study tries to compensate for this challenge and extends our knowledge by testing the assumptions with a more interactive, non-computerized task, i.e., a group discussion, observed by varying panels of observers.

## Study 2

Still, 24% of all leading positions in organizations worldwide are held by females (Grant Thornton International Business Report, 2014). At one point, these female leaders had to go through a selection procedure and demonstrate that they are indeed the best candidate for the job. Interested why some individuals suffer from stereotype threat and others do not, Pinel (1999) argued that a person needs to be aware of a stigma and needs to experience being treated differently because of one's social group. Arguably, people who are more aware of the stereotypes connected to them and who might expect to be stereotyped may entertain more invasive thoughts about whether others judge them on basis of a stereotype instead of their actual performance (Pinel, 2002; Smith, 2004). In contrast, a person who does not expect to be treated or viewed according to a stereotype about his or her social group might not experience such invasive thoughts and will therewith be less distracted from the task.

Indeed, several studies have demonstrated that people systematically differ in their stigma consciousness, i.e., in the extent to which they expect to be stereotyped (Pinel, 1999), and that stigma consciousness influences whether stereotype threat is even experienced (e.g., Brown & Pinel, 2002; Guyll, Madon, Prieto, & Scherr, 2010). Brown and Pinel (2002) found that the more an individual is aware about a stigma concerning his or her social group, the

more he or she expects to be treated stereotypically and is therefore vulnerable to stereotype threat.

Therefore, Study 2 not only aims to replicate findings of Study 1 and to extend results from a single-actor to a more interactive group-based selection procedure (i.e., a classic group discussion as frequently used in assessment centers), but also to study the moderating role of stigma consciousness on the link between gender, transparency, and performance. Similarly to past studies, we assume that stigma consciousness enhances the unequal consequences of transparency on perf ormance. We propose:

*Hypothesis 3:* Stigma consciousness moderates the differential effect of gender and transparency-condition (transparent and non-transparent) on performance. More specifically, the adverse consequences of making the required dimension 'leadership' transparent on women's performance will grow stronger, the higher women score on stigma consciousness.

## Method

**Participants and assessment center exercise.** Seventy-nine advanced university-students and graduates (53 females, *M* age = 26.91, *SD* = 5.04) participated in a one-day assessment center training organized by the university's career center for students and graduates who were currently applying for jobs or were intending to do so in the near future. Among them, 46.9% already held a bachelor's or a vocational degree, 19% a master's degree, 5% a PhD whereas the remaining participants were still completing their bachelor education. Participants' area of study was diverse (e.g., 17% nutritional science, 15% economics, 13% languages).

At the start of the assessment center training (which later also included components irrelevant to the purpose of the current study, such as interviews, pen-and-paper tests, and individualized feedback), participants encountered a 4-person leaderless group discussion ask-

ing them to select a new employee from a list of eight potential candidates. The discussion included components of both a conflict negotiation in that each participant had been given individual criteria that the group's solution had to meet as well as components of a hidden profile (Strasser, 1988) in that relevant information about candidates was distributed across participants. After finishing the group discussion, participants filled out a questionnaire about their stigma consciousness.

**Manipulation and experimental design.** We employed a two (gender: female/male) by two (transparency: non-transparent control-group, transparency on gender stereotyped dimension leadership) between subject design in order to test our hypotheses. At the start of the assessment center, participants were randomly divided into four-person groups which then were randomly assigned to the experimental conditions, thus ensuring equal treatment of all participants within one group-discussion. All participants received preparatory instruction material about the following group discussion. For participants in the transparent leadership condition ($N$ = 31), this material also included the information that the following group discussion aimed to assess their leadership skills. Participants in the non-transparent condition ($N$ = 48) read the instruction manual without learning about the underlying dimension.

**Performance.** Performance was rated by teams of two out of a total of 23 observers (totaling 15 females and 8 males). Pairs of observers differed depending on the assessment day, with observer-teams being homogeneously female (45%), male (12%) or mixed in gender (43%). All observers were blind to the research question, hypotheses, and possible manipulations. Similarly to Study 1, Study 2 employed a validated observation sheet (e.g., König et al., 2007) adjusted for the current study. Again the observation sheet contained examples of desirable and undesirable behaviors representative of leadership as well as blank space for the observer to note down their observations. After undergoing a frame of reference training, observers took notes during the group discussion and subsequently rated all four participants on

their leadership skills, using a 5-point scale from 1 (undesired, unsuccessful behavior) to 5 (desired, successful behavior). The observers' inter-rater reliability (ICC) reached .91.

**Stigma consciousness.** Participants completed a 10-item self-report measure for stigma consciousness (Pinel, 1999). Items such as "Stereotypes about women/men never affected me personally" were phrased according to the participant's own gender and were responded to on a Likert scale, ranking from 1 (strongly disagree) to 5 (strongly agree; $\alpha = .70$).

## Results

As in Study 1, data were analyzed with SPSS Version 21, using regression analyses and PROCESS (Hayes, 2013). In these analyses, participants' leadership performance served as the criterion, gender as the predictor, and the transparency condition, and stigma consciousness as moderators. As expected, results revealed a meaningful albeit only marginally significant three-way interaction between gender, condition and stigma consciousness ($b = 2.08$, explaining an incremental 4% of variance, $F(1, 71) = 3.10$; $p = .08$). More specifically, the interaction effect between transparency and gender depended on participants' level of stigma consciousness, in that this interaction was negligible for participants scoring low ($b = -.20$, $t = -.24$, $p = .80$) and was still non-significant for participants scoring moderate on stigma consciousness ($b = .87$, $t = 1.38$, $p = .17$), but that it did turn significant for participants scoring high on stigma consciousness ($b = 1.93$, $t = 2.04$, $p = .04$). Reversely, gender turned from a non-significant predictor of participants' leadership performance in the non-transparent condition (see Figure 1) to a meaningful predictor in the transparent condition (see Figure 2), reaching marginal significance among participants reporting a moderate ($b = .82$, $t = 1.66$, $p = .10$) and full significance among participants reporting high stigma consciousness ($b = 1.69$, $t = 2.37$, $p = .02$).

Analyzing participants separately by transparency condition further supported this conclusion: In the non-transparent condition, neither gender ($b = 1.04$; $p = .67$) nor stigma

consciousness ($b$ = .02, $p$ = .97) nor their interaction term ($b$ = -.39, $p$ = .67) offered any meaningful prediction of participants' leadership performance ($R^2$ = .006, $p$ = .97). In the transparent condition, however, both gender ($b$ = -3.91, $p$ = .05) and the interaction term ($b$ = 1.70, $p$ = .02) were significant predictors, with stigma consciousness reaching marginal significance ($b$ = -.71, $p$ = .08), thus accounting for a total of 27.5% of variance in participants' leadership performance ($p$ = .03). As can be seen in Figure 2, this effect was only partially due to women performing tentatively, i.e., marginally significantly, worse in the transparent assessment condition, the more stigma consciousness they reported ($b$ = -.71, $p$ = .10). Additionally, the interaction effect was caused by men performing better in the transparent condition, the higher they scored on stigma consciousness ($b$ = .99, $p$ = .04).

**Post-hoc analyses.** Different from Study 1, Study 2 was conducted in a group context with varying gender ratios among both co-participants and observers. Both, a relative underrepresentation of one's own gender (e.g., Inzlicht & Ben-Zeev, 2000) and having a member of the comparison group administer the test (e.g., Marx & Goff, 2005) can reportedly elicit stereotype threat as well, and thus we decided to test for the effect of observers' and participants' gender ratio as possible moderators to the effect of gender and transparency. More specifically, we ran two additional PROCESS models with participants' leadership scores as the criterion, gender as the predictor, and transparency condition as the first moderator. In the first of these models, we added the observers' gender ratio as the second moderator to the equation, in the second model the gender ratio of the other participants. Results revealed that neither observers' gender ratio - either as a main effect ($b$ = .68, $p$ = .24) or as a moderator to either the impact of participants' own gender ($b$ = -.02, $p$ = .21) or to the interaction between gender and transparency condition ($b$ = .03, $p$ = .18) – nor the gender ratio among the other participants (main effect: $b$ = .46, $p$ = .52; moderator-effects: $b$ = -.01, $p$ = .58 on the impact

of participants' own gender, and $b = .01$, $p = .80$ on the interaction between gender and transparency condition) impacted participants' performance. We can therefore conclude that gender of the observer or other participants present did not influence participants' performance.

## Discussion

Study 2 refined the finding reported in Study 1 by showing that transparency not automatically harms performance of participants in the stereotype-sensitive group, but that results also depend on the same participants' conscious awareness of these stereotypes. This awareness moreover did not affect women only, i.e., the members of the stereotype-sensitive group, but affected men, too. This is in line with past research on stereotype lift or stereotype boost, a phenomenon that people experience when they are aware that another group is stereotyped negatively within a domain (e.g., Shih, Pittinsky, & Ambady, 1999; Walton & Cohen, 2003). The boost is likely caused by a reminder of the advantage of one's group compared to the other social group (Smith & Johnson, 2006). In the current case, transparency on the dimension leadership might have functioned as such a reminder to men, particularly men aware of their positive stigma as being the more prototypical leadership group compared to women, possibly leading to higher confidence and thus better performance among men. As a result, transparency again widened the performance gap between men and women, at least for participants who were aware of gender as a source of stereotype. In the end, the route may thus be more complex than originally assumed and supported in Study 1 (i.e., transparency impairing the performance among the stereotype-threatened group), but the result of Study 2 is still comparable: Transparency on the dimension leadership created a relative advantage of male over female candidates that had not been there in the non-transparent condition.

## General Discussion

Together, results of both studies indicate that candidates' benefits from transparent selection procedures are far from universal, thus revising results from earlier research (Kleinmann et al., 1993). More precisely, Study 1 shows that transparency can harm performance depending on the candidates' gender and the dimension revealed. When women learned that the simulation aimed to assess their leadership skills, they performed worse than women who did not have this information. This lower performance may result from the experience of stereotype threat: A transparent dimension that reminds candidates of a stereotype connected to their group may scare them to fulfill that stereotype and may therewith distract candidates from the task at hand, rather than helping them to gain confidence and perform well. Yet, Study 1 also showed that transparency can still boost performance as soon as the dimension revealed has no negative stereotype attached to candidates' social group: When men and women learned that the simulation aimed to assess their planning skills, their performance surpassed that of candidates who did not know the dimension assessed.

Study 2 further supported the role of proposed stereotypes in eliciting the above effects by showing that results also depended on participants' own stigma consciousness: While participants of the threatened group (i.e., women) performed tentatively worse under the transparent condition, the higher they scored on stigma consciousness, men performed significantly better. This also suggests that the effects of transparency are somewhat more complex, however, with quite some variance in reactions among participants of both the stereotypically favored and the stereotypically disadvantaged group.

Combining the results of both studies we can conclude that the dimension revealed may determine whether it harms or benefits the performance of some but not all job applicants. By causing different effects for different groups of job applicants, transparency does not actually provide a fair starting ground for all job candidates, thus casting serious doubt

upon the strongly held assumption that transparency increases fairness between candidates (Gilliand, 1993; Kleinmann et al., 1996). Especially when the task is to select a new employee in a typical male occupation in which most job candidates today are male (e.g., engineering or mechanical position), practitioners should consider carefully whether the performance dimensions should be made transparent, as transparency may quickly undermine the performance of the few female candidates available, not because of lacking skills but because of stereotype threat. At the same time, male participants might experience a boost in performance due to transparency and practitioners would then unintentionally widen the selection gap between genders.

We therefore recommend avoiding making stereotype loaded dimension transparent, especially in occupations that are seen as female or male. If, however, transparency is still desired, practitioners should present those dimension as stereotype free as possible in order to minimize the chances of stereotype threat among one and stereotype boost among another group.

**Direction for future Research**

The two studies presented in the current manuscript are to our knowledge the first studies challenging the assumption that transparency fosters both fairness and performance across all candidates. In doing so, the studies are not without limitations, however. A first limitation, borne out of ethical considerations, is that participants were no true job candidates but participated in the studies in the context of applicant trainings. Given the life-changing stakes involved in actual personnel selection, however, it is well conceivable that actual job candidates' stereotype threat would be enhanced rather than reduced compared to lab settings and low-stakes conditions. Future research should still try to test whether transparency can harm performance under real selection situations, e.g., by employing a within-subject design. If candidates first go through non-transparent selection exercises and then through transparent

exercises, one could not only investigate the effects of transparent dimensions but also determine what further individual differences beside stigma consciousness might play a role. Field studies could also investigate whether the type of exercise influences the extent of stereotype threat. While Studies 1 and 2 employed two different types of selection exercises, both of these exercises present high-fidelity simulations as they are usually used rather late in the selection process. It would be important to know whether other common selection exercises, such as single person presentations, role-plays or even interviews bear the same effect when presented transparently. If for example, research shows that some types of exercises are more prone to cause stereotype threat under transparent conditions compared to other exercises, practitioners could decide more easily when to avoid transparency and when to implement it.

Secondly, the current studies focused on the effect of transparency on one particular social group, gender, while the underlying theory would suggest similar effects across other types of groups (ethnic background, age, etc.). To be able to generalize across dimensions and social groups, also the potential harming effect of transparency with negative stereotypes about men and the potential boosting effect on performance due to positive stereotypes about women should be investigated. For example, König and Eagly (2005) found that men can indeed suffer from stereotype threat when they are tested on social sensitivity, and thus also male candidates might experience stereotype threat when being in a transparent selection procedure for a typical female position (e.g., nursing).

Additionally other social groups could be investigated, such as older and younger employees or other ethnicities. Stereotype threat theory proposes that technically anybody could be victim of stereotype threat as long as a few factors are present: The individual must be aware of the (negative) stereotype, the task must be of importance to the person, and the person knows to be compared to others (Smith, 2004; Steele & Aronson, 1995). Interesting,

though also more complicated, would be the study of more than one social group being activated during personnel selection. Social identity theory suggests that people tend to classify others and themselves into various and different social groups (Tajfel & Turner, 1985). Thus, an older man may categorize himself as belonging to the group of men and also to the group of older employees. Similarly a black woman may self-categorize herself as belonging to the social group of blacks and females. The latter example refers to something known as *double jeopardy or double negative*, categorizing oneself or belonging to two disadvantaged social groups in society (Shorter-Gooden & Washington, 1996). Reversely, the activation of positively stereotyped identities (e.g., being Asian) may also counteract the inhibiting effects associated with negatively stereotyped group identities (e.g., being female) on stereotype sensitive tasks (e.g., Math, Rydell, McConnell, & Beilock, 2009; Shih et al., 1999). Therefore, future research could aim to identity circumstances or cues that need to be present within a personnel selection situation that help individuals to categorize themselves rather to a positive social group than to a social group that might be stigmatized, i.e., aiming to activate a positive self-image.

Finally, results indicate that future research should study affected groups not in isolation of each other but investigate how revealing a dimension that if stereotyped can influence the performance of both the negatively and the positively stereotyped groups, with results likely pointing in opposite directions. If future studies replicate the finding that transparency of one dimension can harm performance of one social group but boost the performance of the other (comparison) group, and extends our knowledge by investigating more social groups and the effects of activating the schema of a certain (positive) social group membership, transparent selection procedures will be understood in more detail and practitioners can weigh the risks and benefits when deciding on fairness considerations.

## Appendix C: Summary of Statistics of Study 3

Table 3.1

*Correlations and descriptive data of Transparency in Personnel Selection Study 1*

| Performance of targeted dimension | *Variable ID* | *V1* | *V2* | *V3* | *M* | *SD* |
|---|---|---|---|---|---|---|
| Gender | *V1* | - | | | | |
| Non-transparent Condition (*n* = 45) | | | | | | |
| Planning Performance | *V2* | .37* | - | | 2.66 | 1.03 |
| Leadership Performance | *V3* | .31* | .62** | - | 3.48 | 1.07 |
| Transparent Condition Planning (*n* = 29) | | | | | | |
| Planning Performance | *V2* | -.14 | - | | 4.02 | 0.62 |
| Leadership Performance | *V3* | .01 | .43* | - | 4.00 | 0.71 |
| Transparent Condition Leadership (*n* = 48) | | | | | | |
| Planning Performance | *V2* | .53** | - | | 2.76 | 1.19 |
| Leadership Performance | *V3* | .57** | .71** | - | 3.09 | 1.00 |

*Note*. Gender is scored as 0 = women, 1 = men; * p < .05; **p < .01 (two-tailed).

Table 3.2

*Correlations and descriptive data of Transparency in Personnel Selection Study 2*

| Performance of targeted dimension | *Variable ID* | *V1* | *V2* | *M* | *SD* |
|---|---|---|---|---|---|
| Gender | *V1* | - | | | |
| Stigma Consciousness | *V2* | -.260* | | 2.79 | .51 |
| Non-transparent Condition (*n* = 48) | | | | | |
| Leadership Performance | *V3* | .010 | -.040 | 2.86 | 1.31 |
| Transparent Condition (*n* = 31) | | | | | |
| Leadership Performance | *V3* | .318 | -.142 | 2.82 | 1.05 |

*Note*. Gender is scored as 0 = women, 1 = men; * $p < .05$ (two-tailed).

## Appendix C.2: Figures of Study 3

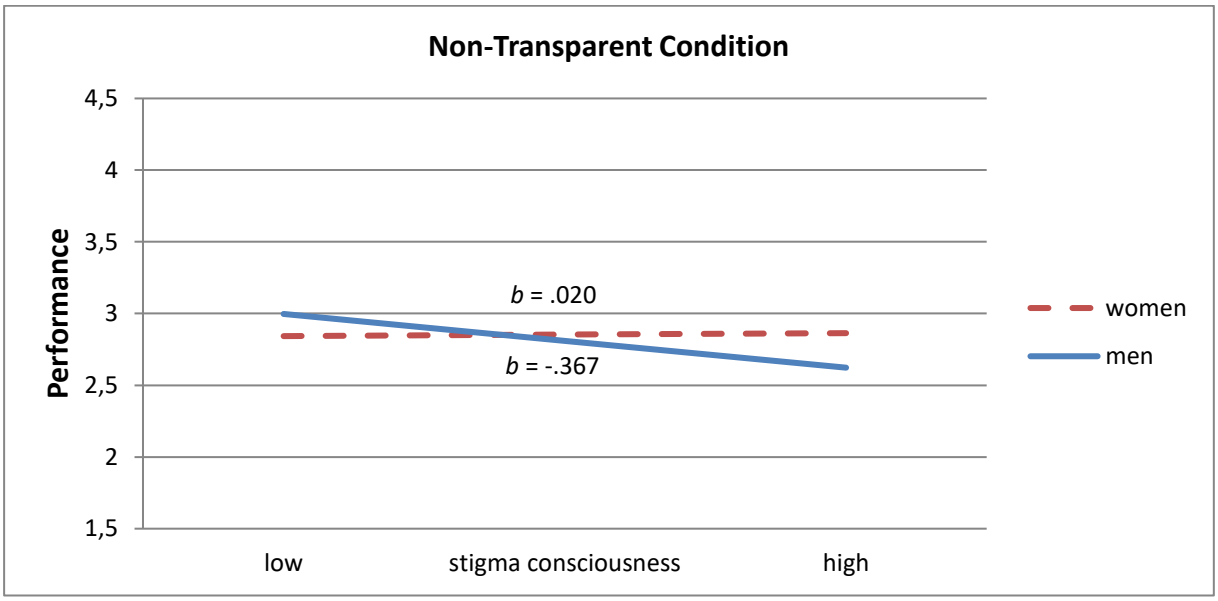


*Figure 3.C1*. Performance and Stigma Consciousness in non-transparent condition

*Note*. Study 2 participants' leadership performance, depending on gender and stigma consciousness in the non-transparent condition.

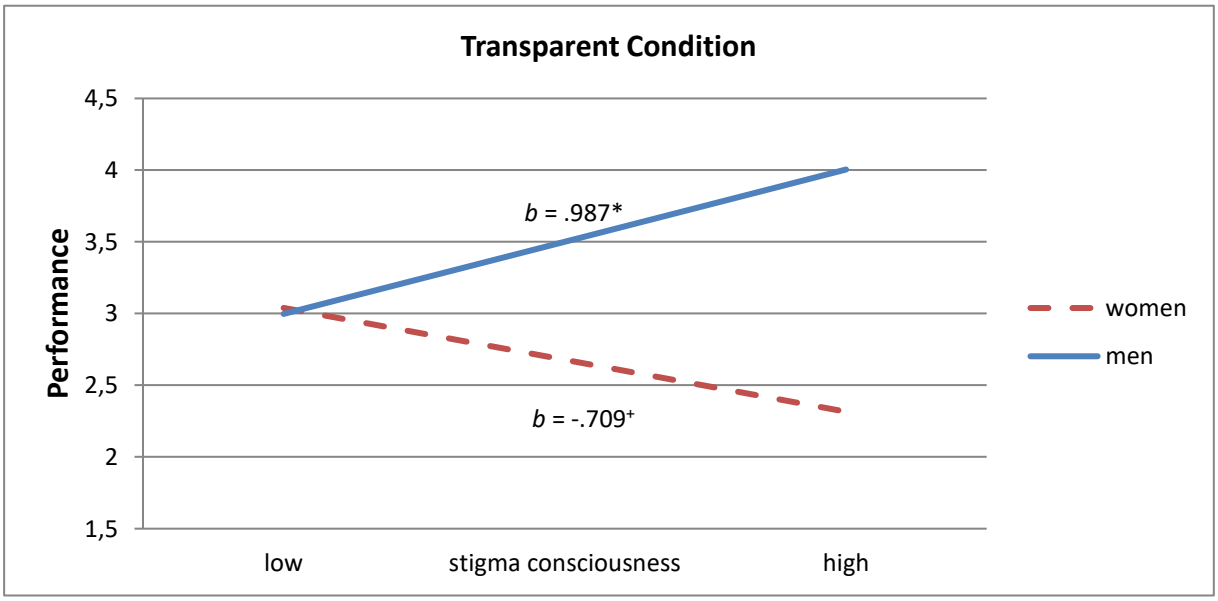


*Figure 3.C2*. Performance and Stigma Consciousness in transparent condition

*Note*. Study 2 participants' leadership performance, depending on gender and stigma consciousness in the transparent condition.

# General Discussion

Leaving the right impression brings along challenges for both the applicant and potential employer. However, both parties have the same interest - the applicant wants the job and the organization wants to fill the position. Even so, on the way to filling the position, organizations and applicants face opportunities and challenges.

Organizations not only want to provide a fair selection situation, but also aim to select the right candidate as objectively as possible. To achieve both fairness and objectivity, organizations choose different selection tools. Regardless of the selection tool implemented or the selection step they are currently in, candidates also try to leave the best possible impression. These two viewpoints often collide and threaten the validity, fairness, and objectivity of personnel selection.

This line of research highlights how the interest of the applicant and the interest of the organization can collide or converge. Further, it shows how applicants attempt to make a favorable impression throughout the selection process, and how some selection tools hinder some applicants' ability to impress positively, while other tools are prone to be influenced by response distortion or IM of applicants.

## Summary of Findings

The first study illustrates that IM behaviors already start when an applicant writes a cover letter and CV. At this point, applicants mostly engage in self-promotion. Research in the interview domain shows that self-promotion can lead to being perceived as more confident and better suited for the job. However, past research on cover letters and CVs shows that it

may also lead to seeming arrogant and over-confident (e.g., Knouse, 1994; Knouse, Giacalone, & Hinda, 1988). Thus, the consequence of IM and their potential influence on decision-making starts before the applicant has the chance to impress in person. Nevertheless, our research shows that IM does not seem to be an individual disposition, but that the selection situation can influence whether or to what extent an applicant shows IM behaviors. Just because an applicant engaged in IM in the cover letter and CV does not mean he or she continues to express IM during the interview. As McFarland et al. (2005) found, the format of the selection instrument seems to influence the extent an applicant engages in IM. Our study underlines this finding by extending our knowledge that cover letters and CVs may elicit IM, but whether IM is shown during the interview depends on the type of interview employed. This is good news for practitioners, because by bearing in mind what effects IM may have, practitioners could decide on a format that minimizes the opportunity to implement IM. Many organizations already use online application formats, i.e., filling in all biographical data on their recruitment page. With such a format and providing specific guidelines on what information should be provided, IM might be minimized. Yet, our line of research does not end with the distorted impression after the first two initial contact points, cover letter/CV and selection interview; it goes further.

Our second study shows that personality tests are also influenced by applicants' response distortion. Again, applicants use the selection method, whether consciously or unconsciously, as an opportunity to present themselves as more favorable than they realistically are. The second study shows that applicants not only distort the personality items of emotional stability and conscientiousness, but also the items of a third, most relevant trait for the specific job they apply to. Past research has already indicated that an extra factor evolves mostly due to distortion on emotional stability and conscientiousness. This factor consists of loading of those two personality traits, but no study to date has shown that other personality items that

are particularly important to a job profile also become part of the extra factor (e.g., Klehe et. al., 2012; Schmitt & Rzan, 1993). Thus, our study replicates the finding of the IEF but extends our knowledge, i.e., shows that the IEF may change according to the specific job requirements. While the first sample of the study helped us to replicate past findings and investigate the loadings onto the IEF between two jobs, the second sample allowed us to compare personality questionnaire answers, taken once under a non-evaluative situation, and once under an evaluative, i.e., simulated personnel selection situation. With the help of this comparison, we were able to illustrate that both the emergence of an additional factor and score inflation essentially represent the same phenomenon. Additionally, we were able to show that the IEF does not only consist of emotional stability and conscientiousness loadings, but that other loadings of traits that are important for the job load highly onto the IEF. The findings of our study therefore make the IEF and/or response distortion more predictable. Practitioners and researchers can either use the emergence of a sixth factor or score inflation as a guide to identify response distortion. Depending on the job, practitioners can even anticipate how the IEF may look, or how applicants might distort theirs answers on personality questionnaires. Next, they can decide whether the correct (i.e., according to the job requirements) response distortion of an applicant is something desirable. As our study also showed, some applicants distort more than others. Past research has found that people with a high understanding of what is expected on the job also perform better during a job interview or even in their jobs (Kleinmann et. al., 2012). Our study indicates that people who correctly identify the personality traits that are required on the job also distort their answers more. However, if one tries to minimize response distortion, the situation can be adjusted. As our study shows, if applicants perceive the test to be difficult, they distort less and in turn, their motivation to distort, yet another predictor we found to relate to response distortion, might decrease. Overall, this study clearly illustrates that response distortion on personality questionnaires is taking place, but we hope to

have shed more light onto how exactly the IEF forms and what other factors might be of influence.

People who engage in IM behaviors in their cover letter or interview, or people who engage in response distortion on personality questionnaires might gain an advantage compared with people who do not engage in such behaviors. Practitioners constantly try to develop selection tools and create environments that minimize the opportunities to engage in such behaviors. By doing so, practitioners and scientists aim to provide a fair starting ground for all applicants and create an environment that helps to focus on the relevant skills and abilities to maximize objective selection.

One situational variable that is used to foster such fairness is transparency. Our last study, however, shows that transparent selection tools do not automatically mean increased fairness for all applicants. The first part of the study indicates that when a dimension includes a stereotype about the applicant's social group, it may influence performance. Past research concluded that providing information about what a selection exercise aims to assess increases the performance of all applicants. However, we showed that this might only be concluded for stereotype-free dimensions. We were able to highlight that providing applicants with the information that planning skills are being assessed, resulted in an increase in all performances. However, when a dimension was revealed that contained a negative stereotype for some applicants but a positive stereotype for others, only the performance of applicants who connect a positive stereotype with the revealed dimension increased. In our study, we therefore showed that revealing information could lead to a performance boost for some and not for others, creating an unfair advantage point. Additionally, we showed that the stigma that a revealed dimension might entail must be conscious to some degree. In the second part of our study, we were able to show that the more aware an applicant was about the social stigma (we argue that such an awareness is influenced by the transparency of the dimension), the more performance

was influenced. We therefore showed that transparency does not automatically cause fairness, but may cause quite the opposite.

Overall, our research aimed to examine, embrace, and discuss several factors that might influence the objectivity of several personnel selection step. However, while our research contains some strength, it also bears some limitations that future researchers could anticipate.

## Limitations and Future Research

With our first study, we were able to show that IM behaviors are indeed expressed in cover letters and CVs, independent of whether the same candidate also expresses IM behaviors during a selection interview. Nevertheless, although our sample size was sufficient for a repeat measurement design, a more diverse or complex sample comparison is needed to gain more understanding on IM behavior during the first contact point. As we showed in our second study, the IEF may change according to the job opening. The same might be true for IM tactics expressed in cover letters and CVs. For example, a person who applies for a legal sector position will engage in different IM tactics than a person who applies for a social sector position. In order to impress, the former might engage in more self-promotion, and the latter in more ingratiation. Additionally, the position a person applies for might also influence the IM tactic chosen. The higher the position, or the more leadership skills are required, the more assertive the IM tactics. As we know rather little about IM tactics in cover letters and CVs and their carried consequences, researchers could try to explore the amount and variety of IM from a real applicant sample, and ideally for different positions and jobs. Our participants underwent a simulated selection procedure, and even though past research has shown that such a simulated setting yields meaningful results (e.g., Kleinmann et. al., 2012; Peeters & Lievens, 2006), an actual applicant sample would not only provide the opportunity to explore IM tac-

tics in a cover letter or CV that was written for a specific job, but could also be used to investigate whether IM tactics influenced decision-making (i.e., who is invited for an interview). Moreover, research could focus on whether IM on a free-written CV or cover letter (i.e., no additional information provided by the hiring organization) is different from the biographical data provided on a recruitment page. Notably, a new trend in personnel selection to optimize the objectivity of applicant screening is to ask applicants to fill out all their information online according to the organization's recruitment page. Often, organizations also clearly instruct applicants how to state their motivation. If IM indeed influences decision-making, it is possible that some formats on a recruitment site minimize the opportunity to express IM behaviors and, in turn, help to control that some applicants engage in it more than others.

The desire to control IM tactics or other forms of response distortion might lead researchers to compare selection formats. Future studies could compare different, but already proven personality questionnaires or answering formats, to investigate whether the formation of an IEF or score inflation is affected. In our second study, in both samples the applicants answered on a scale ranging from low to high agreement. Future research could investigate how the IEF behaves if a forced choice format is employed. Christiansen, Burns and Montgomery (2005) already illustrated that a forced choice format increases validity and decreases response-distortion. However, Christiansen et al. (2005) have used other measures for response distortion, and no study so far has investigated this relationship with the emergence of an IEF or score inflation.

Similarly, the effects of transparency could be explored with such a forced choice format. We argued that transparency of leadership might decrease female performance because the anxiety to conform to a negative stereotype interferes with performance, but what if a female applicant does not have to freely decide or even implement a behavior, but can decide between pre-selected behaviors? A situational judgment test contains such pre-selected behaviors. Having to decide between different options that are offered induces less stress than having to decide freely in a high-pressure situation (e.g., Iyengar, Huberman, Jiang, & Schwarzt,

2000). It is therefore possible that the effects of transparency of stereotype sensitive dimensions are minimized in a situational judgment test. To test the effects of transparency, using a situational judgment test is only the beginning of research in this domain. While we are the first to attempt to illustrate that transparency is not always a performance boost for all applicants, we have so far only investigated this relationship with two different selection exercises. Researchers can focus on a variety of selection instruments to find selection tools that optimize fairness through transparency for all applicants.

## Implications for Organizational Practices

Personnel selection has been under the research lens for several years. Due to such research, organizations have regularly adjusted their selection policies. Our research tries to help organizations evaluate some of the personnel selection tools they might be implementing to a) find the right person for the job, and b) do so fairly.

First, considering that applicants might express IM tactics in cover letters and CVs, practitioners may consider what effects such tactics can have. Our research shows that applicants do engage in IM in their biographical data, but it also shows that the format of the selection tool influences the extent of the IM. IM can be minimized by using a structured interview compared with an unstructured one (Ellis et al., 2002); the same might be true for biographical data. Practitioners could decide to provide greater structure on how the biographical data is presented to the applicant. Future research can determine whether and what kind of structure can minimize the usage and effects of IM in cover letters and CVs. For now, practitioners should keep in mind that cover letters and CVs might impress differently, not only because of different skill levels but also because of IM.

Second, our research helps practitioners not only to understand the IEF better, but also helps them control the effects of something that seems unavoidable and uncontrollable. A vast majority of past research has already showed organizations that the formation of a sixth factor

in personality testing under selection conditions is unavoidable. However, practitioners can now either decide to use the emergence of the IEF or score inflation as a tool to identify applicants who distort more or less, and accurately or inaccurately (i.e., according to specific job demands). Practitioners can even decide to use the IEF as a selection tool. Response distortion has often been viewed as something that carries negative consequences. However, if an applicant is able to correctly identify the personality requirements needed for the job, he or she might also be able to do so successfully on the job. Thus, practitioners may use an applicant's "correct" IEF as criterion to decide whether he or she understands what is expected and needed of him or her on the job. A second selection exercise can then determine whether he/she actually possesses the personality traits that lead to high work performance.

Third, practitioners are most often eager to be fair and objective during personnel selection. However, through our research, practitioners might reconsider when and how they inform applicants about the underlying purpose or dimensions of the testing. By understanding that revealing a dimension that might lead to a performance boost, and potentially, performance loss, practitioners might leave the dimension non-transparent or use an expression that is not a sensitive stereotype. For example, instead of labeling and informing applicants that an exercise is going to assess their leadership skills, practitioners can either not provide this information or choose to label it with a more generic term, such as management skills.

## Concluding Remarks

This dissertation aims to take a different perspective on a variety of topics in personnel selection. These topics are not new in the research domain, but are still in need of more attention. Even though IM has been a popular and well-researched topic, so far the initial first contact between an applicant and hiring organization has been neglected. Response distortion on personality tests can now be understood more clearly, and researchers and practitioners can start to view response distortion as less negative. Transparency might most often lead to a

more fair personnel selection procedure, but now we might consider leaving some dimensions unknown. Overall, this dissertation's objective has been to walk through some of the major selection steps from the perspective of the applicant and the hiring organization, illustrate what we know and what knowledge we lack, and provide greater insight into each of the selection steps. In sum, the picture is still incomplete and in need of many more painters.

## References

Ajzcn, I.(1991). The theory of planned behavior. *Organizational Behavior and Human DecisionProcesses, 50,* 179-211.

Ajzen, I. (2002). Perceived behavioral control, self- Efficacy, locus of ontrol, and the theory of planned behavior. *Journal of Applied Social Psychology, 32,* 665-683.

Ajzen, I. (1985). *From intentions to actions: A theory of planned behavior.* In J. Kuhl & J. Beckman (Eds.), Action- control: From cognition to behavior (pp. 11-39). Heidelberg, Germany: Springer.

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

Amrstrong, P. I., & Fetter Anthoney, S. (2009). Personality facets and RIASEC interests: An integrated model. *Journal of Vocational Behavior, 75,* 346-359.

Strickland, W., Drauden, G., & Martin, C. (1990). Motivational components of test taking. *Personnel Psychology, 43,* 695-716.

Ashforth, B. E., & Mael, F. (1989). Social identity theory and the organization. *Academy of Management Review, 14,* 20-39.

Ashton, M. C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of personality assessment, 91,* 340–345.

Bäckstrom, M., Björklund, F., & Larsson, M. R. (2009). Five- factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality, 43,* 335-334.

Bargh, J. A. (1990). *Auto-motives: Preconscious determinants of social interaction.* In R. M. Sorrentino & E. T. Higgins (Eds.), Handbook of motivation and cognition (pp. 93–130). New York: Guilford Press

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.

Barrick, M. R., Mount, M. K., & Gupta, R. (2003). Meta- analysis of the relationship between the five- factor model of the personality and Holland's occupational types. *Personnel Psychology, 56,* 45-74.

Barrick, M. R., Shaffer, J. A., & DeGrassi, S. W. (2009). What you see may not be what you get: relationships among self-presentation tactics and ratings of interview and job performance. *The Journal of applied psychology, 94,* 1394–1411.

Beaty, J. J., Cleveland, J. N., & Murphy, K. R. (2001). The relationship between personality and contextual performance in "strong" versus "weak" situations, *Human Performance, 14,* 125–148.

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., Cunningham, C. J. L., & Ghorbani, N. (2011). The ubiquity of common method variance: The case of the Big Five. *Journal of Research in* Personality, 45, 417.429.

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.

Bolino, M. C., Kacmar, K. M., Turnley, W. H., & Gilstrap, J. B. (2008). A multi-level review of impression management motives and behaviors. *Journal of Management, 34,* 1080–1109.

Bozeman, D. P., & Kacmar, K. (1997). A cybernetic model of impression management processes in organizations. *Organizational Behavior and Human Decision Processes, 69,* 9–30.

Brown, R. P., & Pinel, E. C. (2002). Stigma on my mind: Individual differences in the experience of stereotype threat. *Journal of Experimental Social Psychology, 39,* 626-633.

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, application, 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.

Chatman, J. A. (1989). Improving interactional organizational research: A model of person-organizational fit. *Journal of Management, 14,* 333–349.

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.

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.

Costa, P. T., & McCrae, R. R. (1995). *The Neo Personality Inventory manual.* Odessa, FL; Psychological Assessment Resources.

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

De Fruyt, F., & Mervielde, I. (1997). The five-factor model of personality and Holland's RIASIC interests types. *Person and Individual Differences, 23,* 87-103.

Dickerson, A., & Taylor, M. A. (2000). Self-limiting behavior in women: Self-esteem and self-efficacy as predictors. *Group & Organization Management, 25*, 191–210.

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

Douglas, C. N. (1917). *Forty Thousand Quotations: Prose and Poetical, comp. by Charles Noel Douglas.* New York: Halcyon House, retrieved from Bartleby.com: http://www.bartleby.com/348/810.html

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.

Ellis, A. P. J., West, B. J., Ryan, A. M., & DeShon, R. P. (2002). The use of impression management tactics in structured interviews: A function of question type? *Journal of Applied Psychology, 87,* 1200–1208.

Elshout, J. J. (1999). De Vijf Persoonlijkheidsfactoren Test (5PFT) 1973- 1999. Nederlandse *Tijdschrift voor de Psychologie, 54,* 195-207.

Fishbach, A., & Trope, Y. (2005). The substitutability of external control and self- control. *Journal of Experimental Social Psychology, 41,* 256-270.

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

Fletcher, C. (1990). The relationships between candidate personality, self-presentation Strategies, and interviewer assessments in selection interviews: An empirical study. *Human Relations, 43,* 739–749.

Folger, R. (1987). *Reformulating the preconditions of resentment: A referent cognitions model.* In J.C. Maters & W.P. Smith (eds). Social comparison, social justice, and relative deprivation: Theoretical, empirical, and policy perspectives (pp.183-215). Hillsdale, NJ:Erlbaum.

Folger, R., & Konovsky, M. A. (1989). Effect of procedural and distributive Justice on reactions to pay raise decisions. *Academy of Management, 32,* 115-130.

Garcia, J. M. R. (2001). Scientia potestas est- Knowledge is power: Francis Bacon to Michel Foucault. *Neohelican, 18,* 109-122.

Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review, 4,* 694–734.

Gilmore, D. C., & Ferris, G. R. (1989). The effects of applicant impression management tactics on interviewer judgments. *Journal of Management, 15,* 557–564.

Goffman, E. (Ed.). (1959). *The presentation of self in everyday life.* Garden City NY: Doubleday.

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

Grant Thornton International Business Report (2014). *Women in business: from classroom to boardroom.* Retrieved from:

http://www.internationalbusinessreport.com/files/IBR2014_WiB_report_FINAL.pdf

Griffith, R. L., & Peterson, M. H. (2006). *A closer examination of applicant faking behavior.* Greenwich, CT: Information Age Publishing.

Griffith, R. L., Chmielowski, T., & Yoshita, Y. (2007). Do applicants fake? An examination of the frequency of applicant faking behavior. *Personnel Review, 36,* 341–355.

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

Guyll, M., Madon, S., Prieto, L., & Scherr, K.C. (2010). The potential roles of self-fulfilling prophecies, stigma consciousness, and stereotype threat in linking Latino/a ethnicity and educational outcome. *Journal of Social Issues, 66,* 113-130.

Hattrup, K., & Jackson, S. E. (Eds.). (1996). *Learning about individual differences by taking situations seriously*. San Francisco: Jossey-Bass.

Hayes, A.F. (2013). *Introduction to mediation, moderation, and conditional process analysis. A regression-based approach.* New York, NY: The Guilford Press.

Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior, 32,* 113–135.

Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology, 92,* 1270-1285.

Hogan, R.T., & Hogan, J.C. (1992). *Manual for the Hogan Personality Inventory.* Tulsa, OK: Hogan Assessment Systems.

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.

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work.* Odessa, FL: Psychological Assessment Resources.

Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology, 75,* 581-595.

Howard, J. L., & Ferris, G. R. (1996). The employment interview context: Social and situational influences on interviewer decisions. *Journal of Applied Social Psychology, 2*6, 112–136.

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.

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 2,* 356–371.

Kacmar, K. M., Delery, J. E., & Ferris, G. R. (1992). Differential effectiveness of applicant impression management tactics on employment interview decisions1. *Journal of Applied Social Psychology, 22,* 1250–1272.

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

Klehe, U-C., Kleinmann, M., Harstein, T., Melchers, K. G., König, C. J., Heslin, P. A., & Lievesn, F. (2012). Responding to personality tests in a selection context: the role of the ability to identify criteria and the ideal-employee factor. *Human Performance, 25,* 273-302.

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., & Klehe, U.-C. (2010). Selling oneself: Construct and criterion-related validity of impression management in structured interviews. *Human Performance, 24,* 29–46.

Kleinmann, M., Exler, C., Kuptsch, C. & Köller, O. (1995). Unabhängigkeit und Beobachtbarkeit von Anforderungsdimensionen im Assessment Center als Moderatoren der Konstruktvalidität. *Zeitschrift für Arbeits- und Organisationspsychologie, 1,* 22-28.

Kleinmann, M., Ingold, P. V., Lievens, F., Jansen, A., Melchers, K. G., & Koenig, C. J. (2011). A different look at why selection procedures work: The role of candidates' ability to identify criteria. *Organizational Psychology Review, 2,* 128-146.

Kleinmann, M., Kuptsch, C., & Köller, O. (1996). Transparency: A necessary requirement for the construct validity of assessment centers. *Applied Psychology: An International Review, 45,* 67-84.

Kling, K. C., Hyde, J. S., Showers, C. J., & Buswell, B. N. (1999). Gender differences in self-esteem: A meta-analysis. *Psychological Bulletin, 125,* 470–500.

Knouse, S. B. (1994). Impressions of the Resume: The Effects of Applicant Education, Experience, and Impression Management. *Journal of Business and Psycholgy, 9(1),* 33–45.

Knouse, S. B., Giacalone, R. A., & Hinda, P. (1988). Impression management in the resume and its cover letter. *Journal of Business and Psycholgy, 3,* 242–249.

Kolk, N. J., Born, M. P., & v. d. Flier, H. (2003). The transparent assessment center: The effects of revealing dimensions to candidates. *Applied Psychology: An International Review, 52,* 648–668.

Kolk, N. J., Born, M. P., & v. d. Flier, H. (2004). A triadic approach to the construct validity of the assessment center. *European Journal of Psychological Assessment, 3,* 149–156.

König, A. M., & Eagly, A. H. (2005). Stereotype threat in men on a test of social sensitivity. *Sex Roles, 7,* 489–496.

König, C. J., Lechers, 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.

Larson, L. M., Rottinghaus, P. J., & Borgen, F. H. (2002). Meta-analyses of big six interests and big five personality factors. *Journal of Vocational Behavior,* 61, 217-239.

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two- component model. *Psychological Bulletin, 107,* 34- 47.

Leary, M. R., & Kowalski, R. M. (1990). Impression Management: A Literature Review and Two-Component Model. *Psychological Bulletin, 107,* 34–47.

Levashina, J., & Campion, M. A. (2006). A Model of Faking Likelihood in the Employment Interview. *International Journal of Selection and Assessment,* 14, 299–316.

Lilienfeld, S.O., Andrews, B. P., Stone-Romero, E. F., & Stone, D. (1994). The relations between self- report honesty test and personality measure in prison and college samples. *Journal of Research in Personality, 28,* 154-169.

Little, T. D., Cunningham, W.A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the Question, Weighing the merits. *Structural Equation Modeling, 9,* 151-173.

Locke, E. A., Mento, A. J., & Katcher, B. L. (1978). The interaction of ability and motivation in performance: An exploration of the meaning of moderators. *Personnel Psychology, 31,* 269-280.

Major, B., Carrington, P.I., & Carnevale, P.J.D. (1984). Physical attractiveness and self-esteem: Attributions for praise from an other-sex evaluator. *Personality and Social Psychology Bulletin, 10,* 43-50.

Marx, D. M., & Goff, P. A. (2005). Clearing the air: The effect of experimenter race on target`s test performance and subjective experience. *British Journal of Social Psychology, 44,* 645-657.

McFarland, L. A., & Ryan, A. M. (2000). Variane in faking across noncongitive 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.

McFarland, L. A., Ryan, A. M., & Kriska, S. D. (2003a). Impression Management Use and Effectiveness Across Assessment Methods. *Journal of Management, 29,* 641–661.

McFaland, L. A., Yun, G., Harold, C. M., Viera, L., & Moore, L. G. (2005). An examination of impression management use and effectiveness across assessment center exercises: The role of competency demands. *Personnel Psychology, 58,* 949–980.

McFarlin, D. B., Sweeney, P. D. (1992). Distributive and procedural justice as predictions of satisfaction with personal and organizational outcomes. *Academy of Management, 35,* 626-537.

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

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. (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.

Mount, M. K., Barrick, M. R., Scullen, S. M., Rounds, J. (2005). Higher- order dimensions of the big five personality traits and the big six vocational interest types. *Personnel Psychology, 58,* 447-478.

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.

O*Net online (2010), retrieved on August 13th, 2010 fromhttp://www.onetonline.org/.

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.

Peeters, H., & Lievens, F. (2006). Verbal and Nonverbal Impression Management Tactics in Behavior Desciption and Situational Interviews. *International Journal of Selection and Assessment, 14,* 206–222.

Peffley, M., & Hurwitz, J. (1998). *White stereotypes of Blacks: Sources and political consequences*. Race and politics survey, retrieved from: http://www.uky.edu/AS/PoliSci/Peffley/pdf/PeffleyHurwitz%201998%20WHITES%20STEREOTYPES%20OF%20BLACKS.pdf

Piaget, J. (1950). *The psychology of intelligence.* London: Keagon Paul, Trench, & Trubner.

Piaget, J., Brown, T., & Thampy, K.J. (1984). *The Equilibration of cognitive structures: The central problem of intellectual development.* Chicago: University of Chicago Press.

Pinel, E. C. (1999). Stigma consciousness: The psychological legacy of social stereotypes. *Journal of Personality and Social Psychology, 76,* 114-128.

Pinel, E. C. (2002). Stigma consciousness in intergroup contexts: The power of conviction. *Journal of Experimental Social Psychology, 38,* 178-185.

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.

Roulin, N., & Krings, F. (2016). When winning is everything: The relationship between competitive worldviews and job applicant faking. *Applied Psychology, 65*, 643–670.

Roulin, N., Bangerter, A., & Levashina, J. (2015). Honest and deceptive impression management in the employment interview: Can it be detected and how does it impact evaluations? *Personnel Psychology, 68*, 395–444.

Roulin, N., Krings, F., & Binggeli, S. (2016). A dynamic model of applicant faking. *Organizational Psychology Review, 6*, 145–170.

Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology, 74*, 629–645.

Rydell, R. J., McConnell, A. R., & Beilock, S. L. (2009). Multiple social identities and stereotype threat: Imbalance, accessibility, and working memory. *Journal of Personality and Social Psychology, 96,* 949-966.

Schlenker, B. R. (1980). *Impression Management*. Monterey, CA: Brooks/Cole.

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review, 2,* 336–356.

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.

Shih, M., Pittinsky, T.L., & Ambady, N. (1999). Stereotype susceptibility: Identity, salience, shifts in quantitative performance. *Psychological Science, 10,* 80-83.

Shorter-Gooden, K., & Washington, N.C. (1996). Young, Black, and female: the challenge of weaving an identity. *Journal of Adolescence, 19,* 465-475.

Smith-Jentsch, K. A. (2007). The impact of making targeted dimensions transparent on relations with typical performance predictors. *Human Performance, 3,* 187–203.

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, J. L. (2004). Understanding the process of stereotype threat: A review of mediational variables and new performance goal directions. *Educational Psychology Review, 16,* 177-206.

Smith, J.L., & Johnson, C.S. (2006). A stereotype boost or choking under pressure? Positive gender stereotypes and men who are low in domain identification. *Basic and Applied Social Psychology, 28,* 51-63.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35,* 4–28.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69,* 797–811.

Steele, C. M., & Aronson, J. (1998). *Stereotype threat and the test performance of academically successful African Americans.* In Jencks, C., & Phillips, M. (Eds.), The Black-White Test Score Gap (pp. 401-427).Washington D.C: The Bookings Institution Press.

Steele, C. M., Spencer, S. J., & Aronson, J. Contending with group image: The psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology, 34,* 379-440.

Steele, C., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69,* 797-811.

Stevens, C. K., & Kristof, A. L. (1995). Making the right impression: A field study of applicant impression management during job interviews. *Journal of Applied Psychology, 80,* 587–606.

Stone, J., & McWhinnie, C. (2008). Evidence that blatant versus subtle stereotype threat cues impact performance through dual processes. *Journal of Experimental Social Psychology, 44,* 445–452.

Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on Black and White athletic performance. *Journal of Personality and Social Psychology, 6,* 1213–1227.

Strasser. G. (1988). Computer simulation as a research tool: The DISCUSS model of group decision. *Journal of Experimental Social Psychology, 24,* 393-422.

Tajfel, H., & Tumer, I. C. (1985) *The social identity theory d intergroup behavior.* In S. Worchel & W. G. Austin (Eds.), Psychology of intergroup relations (pp. 7-24). Chicago: Nelson-Hall.

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.

Tett R. P., Anderson M. G., Ho C.-L., Yang T. S., Huang L., Hanvongse A. (2006). Seven nested questions about faking on personality tests: An overview and interactionist model of item-level response distortion. In Griffith R. L., Peterson M. H. (Eds.), *A closer examination of applicant faking behavior (pp. 43*–83*).* Greenwich, CT: Information Age.

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., McFarland, L. A., & Razmark, P. H. (2007). Antecedents of impression management use and effectiveness in structured interviews. *Journal of Management, 33,* 752–773.

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.

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

Walton, G.M., Cohen, G.L. (2003). Stereotype lift. *Journal of Experimental Social Psychology, 39,* 456-467.

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