

# Faking Good and Personality Assessments of Job applicants: A Review of the Literature

**Dr. Gerry Fahey**  
Principal, OD Solutions  
Dublin, Ireland

© Gerry Fahey. This work is licensed under the Creative Commons Attribution-Non-Commercial-Share Alike 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-sa/4.0/> .

## Abstract

*High stakes selection contexts often drive the provision of socially desirable responses from job applicants. This can take the form of 'faking good' and can lead to inaccurate personality assessments. This article reviews the extant research on the extent to which faking good occurs, the psychological factors that may lead to faking good by job applicants, and how faking good is measured. In particular, the review considers the role of moral hypocrisy in this regard and considers how it can be minimised.*

*Keywords: Personnel management; Personality assessment; Executives--Recruitment*

## Introduction

Since the early 1990's research has shown that in work and organisational settings the dimensions of personality help to explain individual differences in behaviours (Barrick, Mount, and Judge, 2001; Ones, Dilchert, Viswesvaran, and Judge, 2007; Salgado, 2003). The 'Big Five' dimensions have today become synonymous with the topic of personality (Hogan, 2005). This model of personality has its origins in a factor analytic research approach to understanding personality (Digman, 1990). The Big Five dimensions or, alternatively, the Five Factor Model (FFM) of personality have achieved widespread acceptance as a satisfactory explanatory model of the structure of personality and individual differences (Barrick, Mount, and Judge, 2001). The five broad dimensions, or factors, of Neuroticism, Extraversion, Openness, Agreeableness and Conscientiousness have been shown to be replicable across different demographic, ethnic, and cultural groupings (McCrae and Terracciano, 2005; Salgado, Moscovio, and Lado, 2003). According to Salgado (2016), personality inventories are utilised more than biodata, assessment center, and situational interviews in personnel selection. However, concern about the accuracy of self-report personality measures has long been an issue for psychologists (Cronbach and Meehl, 1955). The issue of 'faking good' on self-report personality measures by job applicants is a major concern, according to some researchers (Jeong, Christiansen, Robie, Kung, and Kinney, 2017; Kiefer and Benit, 2016; Roulin and Bourdage, 2017)

In this article the impact of socially desirable responding in the form of faking good by job applicants on the accuracy of personality measures is reviewed. Following this, the role of socially desirable responding in the form of moral hypocrisy as a behavioural manifestation of faking good is considered, along with how research into the role of objective self-awareness can help in dealing with the occurrence of applicant faking good. Finally, how socially desirable responding can also impact on measures of impression management that are used to detect faking good is examined.

### **Personality and Job Performance**

Personality assessment using the self-report measures of the Big Five is now a well-established practice in the field of applied industrial/organisational (I/O), or occupational psychology (Barrick, Mount, and Judge, 2001; Hogan, 2005; Hough and Oswald, 2008; Kiefer and Benit, 2016; Niessen, Meijer, and Tendeiro, 2017; Roberts, Kuncel, Shiner, Caspi, and Goldberg, 2007). Meta-analytic evidence (Barrick and Mount, 1991; Salgado, 2003) suggests that some of the Big Five dimensions are related to overall job performance in virtually all jobs, whereas other dimensions are related to performance in a more limited number of jobs. Conscientiousness has been empirically shown to be a valid predictor of job performance across performance measures in all occupations studied (Salgado, 2003). Neuroticism has also been found to be a generalisable predictor when overall work performance was the criterion, but its relationship to specific performance criteria and occupations was less consistent than Conscientiousness (Barrick, Mount, and Judge, 2001). Extraversion has been found to be related to job performance in occupations where interactions with others form a significant portion of the job such as jobs in the sales and marketing area (Barrick et al., 2001). Agreeableness is a useful predictor of service orientation and teamwork, because it has been demonstrated to have high predictive validity in jobs and work settings that involve considerable interpersonal interaction, particularly when the interaction involves helping, cooperating and nurturing others (Mount, Barrick, and Stewart, 1998). Extraversion and Openness to Experience appear to be related to training proficiency and creativity (Barrick et al., 2001, Salgado, 2005).

Ones, Dilchert, Viswesvaran, and Judge (2007) conducted a detailed review of the most comprehensive meta-analyses that have examined the relationships between the Big Five and the following variables: (a) performance criteria (e.g., overall job performance, objective and task performance, contextual performance, and avoidance of counterproductive behaviours), (b) leadership criteria (emergence, effectiveness, and transformational leadership), (c) other criteria such as team performance and entrepreneurship, and (d) work motivation and attitudes. They showed that the accumulated body of evidence supports the inference that the criterion-related validities of personality measures are substantial. The Big Five personality variables, as a set, do indeed predict important organisational behaviours such as job performance, leadership, and even work attitudes and motivation. The effect sizes for most of these criteria are moderate to strong (Salgado, 2005). Judge, Bono, Ilies, and Gerhardt (2002) have shown that leadership was related to the Big Five dimensions

of personality. Extraversion was found to be the most important trait of leaders and effective leadership. After Extraversion, Conscientiousness and Openness to Experience were the strongest and most consistent correlates of leadership.

Another area of relevance to work and organisational settings in which the study of personality has been fruitful is in the area of career progression and success (Roberts, Kuncel, Shiner, Caspi, and Goldberg, 2007). There are two dimensions to this aspect of an individual employee's life, namely, job satisfaction (intrinsic) and aspects such as salary and position in the organizational hierarchy (extrinsic). Four of the Big Five dimensions have been shown to relate to either extrinsic or intrinsic career success, with Conscientiousness and Extraversion being associated with slightly higher levels of extrinsic and intrinsic career success whilst Neuroticism and Agreeableness are associated with slightly lower levels of career success (Judge, Higgins, Thoresen, and Barrick, 1999). These insights are of use when it comes to career counselling of employees. The effect sizes are small because of the role that moderators, such as family status or industry characteristics, play in determining career outcomes. There are also many contingencies that might alter the relationship between personality and career outcomes (Judge and Kammeyer-Mueller, 2007).

### **Accuracy in Personality Measurement**

Organisations that use personality measures in the selection and assessment of managers, and which retain these employees, are likely to outperform their competitors who do not select on the basis of personality (Hogan, Hogan, and Kaiser, 2010; Oh, Kim, and Van Iddekinge, 2015). These benefits of personality assessment, however, only arise if the individual's true score on a personality measure is accurately assessed, thus being indicative of 'construct validity'. 'High stakes' contexts occur when the assessment can play a determining role in deciding who will gain access to an employment opportunity (Sackett, Schmitt, Ellingson, and Kabin, 2001; Ellingson, Heggstad, and Makarius, 2012). The reward available is, to some extent, dependent on the outcome of the assessment (Wise and Demars, 2005). On the other hand, 'low stakes' assessment situations are those in which participants are assessed in settings that do not have a potential for a similar reward. Research linking personality and job performance, using job incumbents as participants, is in this category. In defending the use of personality measures in applied contexts such as high stakes employee selection situations, Hogan (2005) stated "the problem is that business people have trouble getting good advice from academic psychology. This, in turn, explains the widespread interest in bogus measures of personality such as the Myers–Briggs Type Indicator and Goleman's Emotional Competence Inventory" (p. 334).

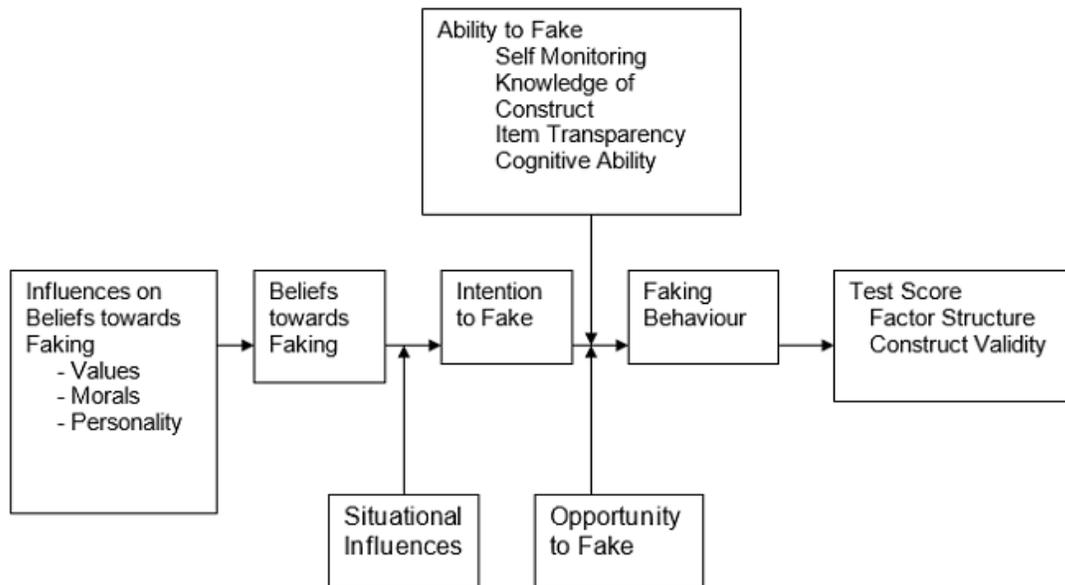
The accurate assessment of personality in selection contexts is an important theoretical question (Ellingson, 2012; Ployhart, Schmitt, and Tippins, 2017), with critical consequences in applied settings (Griffith and Converse, 2012). Objectively scored psychometric measures, such as cognitive ability tests, differ from personality measures which rely on self-reported data. The items in ability measures are

objectively scored, whereas there is the possibility of inaccuracy due to faking good by the individual being assessed in self-report personality measures.

An individual's test score has meaning only on the basis that it provides a measurable link with the individual's behaviour in a particular setting of interest (Lance, Dawson, Birkelbach, and Hoffman, 2010). However, the question of whether individuals respond honestly to the items in personality measures, or whether they engage in what is usually referred to either as 'faking good' or 'impression management' is a hotly contested topic (Ellingson, 2012). Some researchers maintain that faking good is a serious problem because of inaccuracy resulting from faking good (Griffith and Converse, 2012; Jeong et al., 2017; Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt, 2007). Other researchers argue that, even if faking good were to occur, it would not matter because studies have shown that the occurrence does not affect the criterion-related validity of self-report personality measures (Ones, Viswesvaran, and Reiss, 1996). A few researchers maintain that faking is not an issue in high stakes personality assessments (Hogan, Barrett, and Hogan, 2007). These opposing viewpoints raise issues pertaining to core psychometrics that question the construct validity of personality measures. This is because they raise doubts about the accuracy of the inferences made about an individual based on her or his score on a personality measure in a personality assessment context, particularly those that are described as 'high stakes'. The construct validity of personality measures in selection situations is a fundamental measurement issue with respect to actual inferences that are made about the personality traits of job candidates.

Figure 1 contains a suggested nomological network (McFarland and Ryan, 2000) for the various factors that affect a test taker's observed score on a personality measure. Knowledge of the nomological network is essential for an understanding of the construct validity of a psychological measure (Cronbach and Meehl, 1955; Messick, 1995). If the observed score on such measures is not aligned with the individual's true score on a putative latent construct of each of the dimensions of personality, then the observed score is not a valid measure and is as open to the same criticism as that levelled by Hogan (2005) at the Myers-Briggs Type Indicator or Goleman's Emotional Competence Inventory. Both of these measures lack construct validity and have poor psychometric properties, apart from the MSC EI measure (Bess and Harvey, 2002; Mayer, Salovey, and Caruso, (2008).

**Figure 1 - Nomological Network for Test Scores on a Personality Measure**



The importance of accurate personality assessment of job applicants lies in its value in aiding a better understanding of human behaviour in work and organisational settings both at an individual and aggregated level. Personality predicts job performance, but it is not the only predictor (Hogan, Hogan, and Roberts, 1996). A meta-analysis of personality and overall assessment centre ratings (OAR's) conducted by Collins et al. (2003) found that, although cognitive ability alone predicted much of the variance in OAR's, the addition of personality traits to the model significantly increased the variance accounted for. In fact, they showed that in certain contexts, the combination of a set of personality traits and cognitive ability can predict nearly all of the variance in performance ratings.

## **Socially Desirable Responding**

Socially desirable responding is typically defined as the tendency to give positive self-descriptions (Ziegler, MacCann, and Roberts, 2012), but it can also manifest itself in clinical settings as a tendency to give negative self-descriptions (Perinelli and Gremigni, 2016; Salgado, 2016; Sollman and Berry, 2011). Paulhus (1984) showed that the person being assessed may either be consciously engaging in a deliberate strategy of misrepresentation to make an impression on those who might eventually see his or her personality profile, or the misrepresentation could occur at an unconscious level and be motivated by a latent need for self-enhancement and ego maintenance.

Socially desirable responding can therefore present a problem when accuracy in the assessment of personality is a concern, and, from a construct validity perspective, it must be taken into account (Salgado, 2016). This problem with self-report measures is an old one, and the link with behaviour was pointed out a long time ago by La Piere (1934) in his classic paper on the relationship between attitude and behaviour in which he stated, "Yet it would seem far more worthwhile to make a shrewd guess regarding that which is essential than to accurately measure that which is likely to prove quite irrelevant" (p. 237). In spite of La Piere's admonition, accurate measurement can still be a problem in the personality assessments of job applicants. Socially desirable responding is more likely to occur in situations where there is a desirable outcome at stake, such as in high stakes selection contexts (Griffith, Chmielowski, and Yoshita, 2007), that can affect the rank order of individual job candidates (Morgeson et al., 2007). This was one of the reasons that Campbell and Fiske (1959) advocated the use of their multi-trait, multi-method (MTMM) approach to determining psychometric test accuracy. They were concerned with "the adequacy of tests as measures of constructs, rather than the adequacy of a construct as determined by the confirmation of theoretically predicted associations with measures of other constructs" (p. 100). Socially desirable responding in the form of applicant faking good should therefore be a major concern with regard to the measurement of true scores in personality psychology (Backstrom, Björklund, and Larsson, 2009; Bangerter, Roulin, and König, 2012; Chan, 2009; Morgeson et al., 2007). For example, the transparency of the item content in an omnibus personality measure is one measurement issue that can lead to an occurrence of this problem. Because of item transparency, test takers can develop reasonably accurate hypotheses about what trait an item is tapping into (Schmit and Ryan, 1993).

Today, there appears to be a dominant consensus view about faking in the literature, at least in occupational settings, that:

1. Faking can and does occur (Birkeland, Manson, Kisamore, Brannick, and Smith, 2006; Griffith and Converse, 2012; Landers, Sackett, and Tuzinski, 2011; Markus, 2006, Salgado, 2016)
2. It does not affect the criterion-related reliability of personality measures with respect to research, at the aggregate level, on job performance (Hogan et al. 2007; Li and Bagger 2006; Ones et al. 1996), although Salgado (2016) disputes this. This is because this research relies on samples of job incumbents.
3. It is more likely to occur in 'high stakes' situations such as recruitment where there is a desirable outcome at stake - getting a job (Ellingson, 2012; Griffith, Chmielowski, and Yoshita, 2007).
4. It can have a negative impact on the rank ordering of individual job applicants when personality measures are used in selection contexts (Griffith and Converse, 2012; Hollenbeck, 2009).

Estimates from extant research as to the extent of the occurrence of faking good vary, as well as the extent to which each of the Big Five is prone to being faked. For example, Hogan et al. (2007) are of the view that 'all (applicants) faking all the time'

(p.1280), meaning that the issue of faking good is a non-issue in their view. This strong view has been challenged by several researchers (Sackett, 2012). Research findings on the link between job performance and personality traits using job incumbents (those in employment), is not affected by faking good (Griffith and Converse, 2012). Arthur, Glaze, Villado, and Taylor (2010) demonstrated that, in the aggregate, all of the Big Five dimensions of personality are prone to faking good among job applicants. They also found that, although most test takers' scores were stable, fairly sizeable percentages of the test takers displayed evidence of higher scores as job applicants, compared to their scores later on when they were job incumbents. According to Ziegler et al. (2012),

Faking represents a response set aimed at providing a portrayal of the self that helps a person to achieve personal goals. Faking occurs when this response set is activated by situational demands and person characteristics to produce systematic differences in test scores that are not due to the attribute of interest (p. 8).

As such, it is a form of 'moral hypocrisy'. There is evidence from research (Batson, 2008) that there are two psychological processes involved in socially desirable responding manifested in the form of faking good. These are moral hypocrisy and objective self-awareness (Duval and Lalwani, 1999).

### **Moral hypocrisy**

In two landmark series of experimental psychology studies, Batson and colleagues (Batson, Kobrynowicz, Dinnerstein, Kempf, and Wilson, 1997; Batson Thompson, Seufferling, Whitney, and Strongman, 1999) introduced the concept of moral hypocrisy which they defined as follows: "*moral hypocrisy*: Morality is extolled - even enacted - not with an eye to producing a good and right outcome but in order to appear moral yet still benefit oneself" (p. 1335). People who sincerely value morality, and who firmly believe that they should not put their own rights and interests ahead of the parallel rights and interests of others, can act in ways that seem to show a blatant disregard for the moral principles they hold dear (Batson et al., 1999). The difference between moral disengagement and moral hypocrisy is to be found in the second part of Batson et al.'s (1997) definition. Unlike moral hypocrisy, those who engage in moral disengagement are not necessarily attempting to appear moral while simultaneously engaging in behaviour that is morally lacking. Moral disengagement therefore occurs when individuals deliberately disengage from self-censorship. When moral hypocrisy occurs there is a tangible benefit gained by the individual.

Batson, et al. (1997) presented participants with a moral dilemma. Participants were required to assign themselves and another fictitious participant, to either of two tasks that are different (Batson, 2008). Participants were led to believe that the other (fictitious) participant would not know that they – the actual subjects of the experiment - were allowed to assign the task. One task allowed the participant the chance to earn a raffle ticket and had positive consequences in that it was an enjoyable task. The other task had no chance to earn a raffle ticket and was described in the briefing of

participants as dull and boring. Most participants – (up to 80%) – assigned themselves to the more interesting and rewarding task. Yet only 10% of participants believed that assigning the dull and boring task to the other (fictitious) participant was the moral thing to do.

There are strong parallels between Batson et al.'s (1997, 1999) experiments and the applied situation faced by individual job applicants (Ellingson et al., 2012; Fan et al., 2012; Fischbacher and Föllmi-Heusi, 2013). Completing self-report personality measures in a selection situation is a win/lose situation with an applicant who fakes good possibly winning if the applicant with the 'best' true scores does not fake good or lie. Therefore, the opportunity for moral hypocrisy is present in faking good situations. The job applicant can deliberately lie which is immoral behaviour regardless of the euphemistic labelling of the lying as faking good or impression management. The results of the Batson et al.'s series of experiments provide support from experimental psychology for the level of faking good that Griffith and Converse (2012) maintain are the norm in applied settings – 30% of applicants fake good with a margin of error of + or – 10%.

Batson, et al. (1999) conducted a further experiment to examine whether moral standards being made salient, prior to the opportunity to behave, would actually affect the behaviour of the participants. They tested four experimental conditions. Awareness of moral standards could be either high or low. The salience of moral standards could also be either high or low. This resulted in a 2x2 experiment i.e. 4 possible experimental conditions. They found that participants in the low-standard-salience/high-self-awareness condition responded very differently to those in the high-standard-salience/high-self-awareness condition. In the latter condition, the majority of participants agreed that the most moral way to assign the tasks was to give the positive consequences task to the other participant, whereas in the low-standard-salience/high-self-awareness condition, only a small minority agreed with this. This shows that it is not unreasonable to expect that in the context of high stakes personnel selection, faking good, a form of moral hypocrisy, can be greatly reduced by following a procedure that mimics the procedures that Batson et al. (1997) followed in their experiments (Ellingson et al., 2012; Fan et al., 2012). In recent years, a number of other researchers, including Mazar, Amir, and Ariely (2008) and Shu, Gino, and Bazerman (2011) have expanded on the work of Batson and colleagues and have found similar results.

The importance of moral hypocrisy in faking good is of relevance when the validity of a personality measure is being discussed (Fahey, 2017). It should be borne in mind that aggregated anonymous personality assessments are used for the purposes of establishing the concurrent and/or predictive criterion-related validity of the Big Five dimensions of personality. As mentioned earlier, criterionrelated validation studies usually use job incumbents for their research. This differs with respect to the aspect of construct validity of the same Big Five measures that applies when they are used for the purposes of selecting the 'winner' in high stake job selection situations (Embretson, 1983; Messick, 1995; Sackett, 2012). This aspect of construct validity is

described by Messick as the 'consequential' aspect. There is an economic payoff to be gained in the latter situation – either a promotion with higher status and remuneration, a change of employer to a more putatively desirable one, entry into the workforce, etc. All of these desired outcomes for the applicant have a potential payoff attached that is frequently monetary.

It is arguable that the upper limit of 40% from Griffith and Converse's research is supported by the research on moral hypocrisy (Mazar et al., 2008; Shu et al., 2011). The extant research has shown that lying, cheating, or moral disengagement are heterogeneous phenomena that are situationally dependent (Figure 1). In any given situation where ambiguity is present, moral hypocrisy can occur even though some individuals don't lie or cheat at all, while some others do lie or cheat to the maximum extent possible, and the remainder may lie or cheat to a smaller extent. Executive selection is a form of high stakes selection in which the outcome is frequently that of selecting a preferred candidate from a short list or subset of candidates. In essence, the selection decision rests on a rank ordering, formal or otherwise, of the short-listed candidates. If faking good occurs with some participants in spite of the effectiveness of procedural precautions taken to prevent it, and if it can be measured, the construct validity of the personality measures can still be questioned. This is because of the consequential aspect (Messick, 1995). Even though the success criteria for these types of positions can be difficult to define (Highhouse, 1998; 2002), if the selection criterion or criteria includes personality dimensions then the rank order of candidates can easily change. If even one of the candidates deliberately engages in faking good, the rank order can change when compared to a ranking based on all of the candidates' true scores on these dimensions (Komar, Brown, Komar, and Robie, 2008; Rosse, Stecher, Miller, and Levin, 1998). The consequences of a failure to achieve the objective of accuracy in personality measurement in high stakes employee selection contexts, due to moral hypocrisy in the form of faking good, can be severe in situations such as high stakes selection contexts because of rank order effects. A disproportionate likelihood of those who fake good being selected, which can easily arise from unfairness in testing, has been found consistently in research examining the effect of faking on the rank-order of those selected (Ellingson et al., 2001; Hough, 1998; Komar et al., 2008).

### **Objective self-awareness**

For objective self-awareness to occur, an individual's attention has to be directed inward, and the individual's consciousness must then be focused on himself or herself. Batson et al. (1999) found that by manipulating objective self-awareness it was possible to eliminate moral hypocrisy. The coin toss condition in the Batson et al. (1997, 1999) series of experiments allowed for ambiguity in that participants were able to pretend that the task assignment decision depended on the outcome of a coin toss. All participants had to do in this experiment was to say that they had used the coin toss outcome, which was not independently verified, in making their choice. This ambiguity allowed participants to appear to behave in a moral manner while not being prepared to pay the cost of so doing. The result of the elimination of the coin toss

option resulted in participants in the low-standard-salience/high-self-awareness condition responding very differently to those in the high-standard-salience/high-self-awareness condition. In the latter condition, the majority of participants agreed that the most moral way to assign the tasks was to give the positive consequences task to the other participant, whereas in the low-standard-salience/high-self-awareness condition only a small minority agreed with this. The elimination of the coin toss option removed ambiguity from the experiment which in turn reduced the incidence of moral hypocrisy.

This finding shows that it is not unreasonable to expect that in the context of applied high stakes personnel selection, faking good, a form of moral hypocrisy, can be greatly reduced by following a procedure that mimics the procedures that Batson et al. followed in their experiments. Therefore, the importance of the Batson et al. (1997) and Batson et al. (1999) experiments, lies primarily in the fact that their research findings arguably support the use of a procedural measure that makes moral standards salient for the participants. The statement *“Most participants feel that giving both people an equal chance – by, for example, flipping a coin – is the fairest way to assign themselves and the other participant to the tasks”* was used in the Batson et al. (1997) research to make moral standards salient for participants. In addition, the Batson et al. (1999) experiments arguably showed that a combination of making moral standards salient together with heightened self-awareness eliminated moral hypocrisy. In a later series of experiments, moral hypocrisy in the form of dishonesty was eliminated when the experimenter’s objective scoring of the answers to questions on a test was compared with the participant’s self-scoring of the answers to the test questions (Mazar et al., 2008).

Duval and Wicklund’s theory of self-awareness (Duval and Lalwani, 1999) posits that focussing one’s attention on the self induces a state of objective self-awareness and this leads to an awareness of the discrepancies between the ideal and actual self (Higgins, 1987). For objective self-awareness to manifest itself, the individual engages in introspection and self-evaluation while, at same time, ignoring endogenous environmental factors (Silvia and Duval, 2001). Self-awareness can be experimentally induced by exposing participants to self-focusing stimuli (Morin, 2011). The important aspect of ‘objective self-awareness’ theory in understanding the phenomenon of faking good is the degree to which a person’s attention is focused upon a salient within-self discrepancy (e.g. perceived self-evaluated discrepancy between actual and ideal or ought self in high stakes personality assessment). Provided that attention cannot be directed elsewhere, and not on moral values or standards because of low saliency, there will be efforts to reduce that discrepancy between the actual and ideal or ought self by faking good (Pryor, Gibbons, Wicklund, Fazio, and Hood, 1977). An individual completing a self-report personality measure is focussed on the self, by definition. Based on Batson et al.’s (1999) research, this focus on, and real time awareness of, the discrepancies between the actual and ideal self in a setting such as that of a high

stakes selection situation was shown to be conducive to a state of moral hypocrisy among participants in the a low moral standards saliency context .

The low-standard-salience/high-self-awareness condition is commonplace in life, according to Batson et al. (1999), a circumstance they described as 'frightening'. Their concern arises from the fact that people are frequently asked to make moral decisions in circumstances in which the relevant moral standards are not stated in advance, or when others are not watching, or when their actions are not challenged, or when they do not actually feel accountable for their actions, and so on. Therefore, according to Batson et al. (1999), many everyday moral decisions occur in low-standard-salience/high-self-awareness situations. High stakes employee selection contexts would be a good example of such moral decision making situations. This confirms the relevance and importance of the Batson, et al., experimental findings to the issue of applicant faking good, (1997, 1999). Individuals who fake good, when completing self-report personality measures in high stakes selection situations, could and probably would feel that they acted morally if questioned after completing the questionnaire, even though they actually acted in a manner that served their self-interest. Individuals who are candidates for jobs are acting from a self-interest perspective. They are seeking gains such as better employment and career opportunities, increased income, or some other such long and short term economic benefit. In addition, they are in competition with others for the positions on offer. To at least minimise moral hypocrisy in high stakes selection contexts the assessment situation should be one that provides for both high self-awareness and high moral salience.

There have been a number of studies by other researchers into the question of moral hypocrisy and objective self-awareness (Mazar, Amir, and Ariely, 2008; Shu, Gino, and Bazerman, 2011; Shu, Mazar, Gino, Ariely, and Bazerman, 2012). Mazar, et al. (2008) investigated the extent to which people who think highly of themselves in terms of honesty make use of various mechanisms that allow them to engage in dishonest behaviour while retaining positive views of themselves. Even though participants knew that they were over-claiming their actions it did not affect their self-concept in terms of honesty. By making moral codes salient, cheating was eliminated. Shu et al. (2011) found that having participants read an honour code reduced cheating by half in the Mazar, et al., (2008) experimental design. When participants read and signed the honour code, cheating was eliminated. Behavioural economists have also investigated lying and honesty in a range of experimental settings. Fischbacher and Föllmi-Heusi (2013) developed an experimental design that makes it possible to detect lies when participants face no threat of being caught individually. Their research showed that some participants lied to the fullest extent possible, while some were fully honest. Other participants also lied, but not to the maximum possible extent (Fischbacher and Heusi, Pruckner and Sausgruber, 2013). This research also shows that individuals cheat and lie in permissive situations where there is an opportunity for self-interest gain, just as the Batson et al. (1997, 1999) experiments demonstrated. Once people begin to behave dishonestly by cheating, they disengage morally. On the other hand, it is relatively easy to prevent this moral disengagement and moral hypocrisy by simple

environmental nudges. However, even with assessment conditions of high self-awareness and high moral salience, applicant faking good may still occur. For this reason the assessment procedures need to include a measure that is capable of detecting faking good. Typically, this is done through the use of impression management measures that are sometimes referred to as lie scales.

### **Lie and related scales**

Lie scales are self-report measures that are sometimes used to try to detect and measure socially desirable responding in personality assessments. According to MacCann et al. (2012), these scales are frequently used in applied settings such as personality assessments. The construct validity of these measures has been hotly debated for many years (Block, 2010; Burns and Christiansen, 2011; Connelly and Chang, 2016; Costa and McCrae, 1992; Dilchert and Ones, 2012; MacCann, et al., 2012; Paulhus, 1984). In fact, the use of lie scales of any form to detect faking in self-report measures has been questioned by some researchers, some of whom even go so far as to recommend against using lie scales as a method for detecting faking good (MacCann et al., 2012). There are two categories of such measures in use – unidimensional lie scales which assume that there is a single latent factor underlying the lie scale, and scales based on two dimensions (Paulhus, 1984). The unidimensional scales that are in use can be either stand-alone measures, or, they may be embedded in the personality measure.

In 1930, Hartshorne and May developed a lie scale to detect and help deal with socially desirable responding. High scores on the lie scale were assumed to be indicative of a dishonest character (Paulhus, 2002). Later on, in clinical settings, the widely used omnibus Minnesota Multiphasic Inventory (MMPI) included an embedded socially desirable responding scale, the MMPI Lie Scale, designed to identify individuals deliberately dissembling their clinical symptoms (Hathaway and McKinler, 1989). The Eysenck Personality Inventory (EPQ) also contained a lie scale (Eysenck, 1968). Additionally, two widely used stand-alone socially desirable responding detection scales are also in use (Paulhus, 2002), namely, the Marlowe Crowne SD Scale (MCSD) and the Edwards SD Scale (ESD). These latter two measures contained items claiming improbable virtues and denying common human frailties. The items used in these measures are those with self-descriptions that were not just positive, but improbably positive. Typical items include the following - “I always try to practice what I preach” from the MCSD, and “No one cares much what happens to you” from the ESD (Shaver, Brennan, Robinson, Shaver, and Wrightsman, 1991). A number of the omnibus personality measures in use today such as for example Eysenck’s EPQ and the 16PF (Conn and Rieke, 1994) have the lie scale embedded in the personality measure (Ellingson et al., 2001). These embedded measures have been the most widely employed applied technique to deal with applicant faking (Barrick and Mount,

1996; Holden, 2008; Hough, 1998; Hough, Ones, and Viswesvaran, 1998; Kurtz, Tarquini, and Lobst, 2008).

It is claimed that the scores on these measures have been used by some researchers and practitioners to partial out the variance in personality responses associated with faking in an attempt to obtain more accurate estimates of the construct validity of personality tests (Smith and Ellingson, 2002). The scores on the personality measure are adjusted using the lie scale scores (Dilchert and Ones, 2012). However, this approach has little empirical support as a valid technique for eliminating common method variance (CMV) due to faking (Dilchert and Ones, 2012; MacCann et al., 2012). Moreover, the unidimensional nature of lie scales was questioned by a number of researchers as far back as the 1960's, which led to Paulhus examining the factor structure of the lie scales in use (Paulhus, 2002).

The factor analysis studies of Paulhus in the 1980's of the various socially desirable responding measures that were in use at that time found that there were two, rather than one, socially desirable responding factors. These he referred to as self-deception enhancement and impression management, respectively (Paulhus, 1998). Self-deceptive enhancement (SDE) refers to an unconscious positive bias in item responses. It might occur when individuals complete the self-report measures with the aim of protecting positive self-esteem. In contrast, impression management (IM) refers to the conscious dissimulation of item responses with the aim of making a favourable impression on others (Paulhus, 2002). Sackett's (2012) multiple component analysis of systematic variance in responding to items in a self-report measure is consistent with this latent two factor structure of socially desirable responding of Paulhus (1984). Sackett (2012) differentiated between what he termed 'erroneous self-perception' and 'situationally specific intention' (p. 331) distortion. The former is an automatic response mode whereas the latter is a controlled response mode. In the automatic response mode the test taker has a tendency to automatically respond to items in terms of her or his best self. Faking good is deliberate (Griffith and Converse, 2012) and is related to IM unlike SDE, whereas item response distortion arising from SDE is not deliberate (Ellingson, 2012; Lönnqvist, Irlenbusch, and Walkowitz, 2014).

Moral hypocrites value morality and can firmly believe that they should not put their own rights and interests ahead of the parallel rights and interests of others (Mazar et al., 2008). Yet, in spite of this, they can act in ways that seem to show a blatant disregard for the moral principles they hold dear (Batson et al., 1997). Faking good is a response set to questions in a self-report personality measure that is intended to provide a false portrayal of the self that helps the individual achieve personal goals and, as such, is a form of moral hypocrisy (Ziegler et al., 2012). Faking good and impression management are forms of socially desirable responding that are not always easy to detect. This form of social desirability responding is a prominent concern in applied, high-stakes, assessments that use personality traits (Fahey, 2017). Some of the inferences made based on an individual candidate's scores on a personality test may still not be valid in spite of measures used in the assessment process to prevent

faking good, a manifestation of moral hypocrisy, by job candidates from occurring. Lie scales can be a useful tool in identifying such 'fakers', provided they are accurate and valid measures (Connelly and Chang, 2017). Connelly and Chang emphasised that understanding and measuring the negative influence of exaggerated responding should remain a focal concern for personality and applied psychologists.

As a result of his research, Paulhus (1984, 1998) developed a forty item measure – the Balanced Inventory of Desirable Responding (BIDR) - to measure the two factors of socially desirable responding. All of the forty items are affirmation statements, and there are equal numbers of attribution and denial items for each of the two 20-item sub scales measuring Self-Deceptive Enhancement (SDE) and Impression Management (IM). The BIDR is the most widely used standalone socially desirable responding measure in both research and applied settings (Ellingson, Heggstad, and Makarius, 2012). The IM scale of the BIDR has been widely used in the study of the effect of faking good on personality measures in occupational settings (Ellingson et al., 2012; Fan et al. 2012). This scale has also been used, as a method for measuring impression management, in research into the hierarchical structure of personality (DeYoung, Peterson, and Higgins, 2002; Ellingson, 2012; Fan et al., 2012).

Arguably, Paulhus's (1984) distinction between the two dimensions of socially desirable responding is the more construct valid measure of impression management because of the volitional nature of IM compared with SDE. The other measures of socially desirable responding, such as the unidimensional lie scales, do not distinguish between these two dimensions, (Lönqvist et al., 2014; Sackett, 2012). The findings of Lönqvist et al. (2014) support the contention that the BIDR-IM scale does indeed assess the tendency to consciously, rather than subconsciously, give inflated self-descriptions to an audience. The extant research would also suggest that the BIDR-IM scale measures faking good because it is a form of moral hypocrisy manifested as intentional distortion similar to that observed in the research studies of Mazar, Amir, and Ariely (2008), and those of Shu, Gino, and Bazerman (2011).

### **Item transparency in impression management measures**

Procedural controls, according to Podsakoff, MacKenzie, and Podsakoff (2012), need to be in place in order to deal with the problem of item transparency, which can lead to faking good because of the clear-cut nature of the items being grouped together. The items in the BIDR-IM scale are, deliberately, both overt and clear cut. The research findings, reviewed earlier, concerning moral hypocrisy would suggest that this effect can also be manifested in impression management measures in the context of the assessments of participants in high stakes selection contexts. This presents a construct validity concern (Embretson, 2007) in using impression management measures in high stakes selection situations – presenting the twenty items together as a group to job applicants in order to detect 'fakers' in selection situations is a concern arising from the overt and clear-cut nature of the items in the self-report measures (Podsakoff et al., 2012). Based on the research of McFarland, Ryan and Ellis (2002),

into the effect of random item placement compared to item grouping in self-report questionnaires, job applicants are likely to engage in faking good. It is argued that they would be likely to recognize the objective of the IM items if they were presented as a single group of 20 items, as is the case with the BIDR measure. The psychometric properties of the personality measure used by McFarland et al. (2002) were found to be better when the items that measure the same construct were randomly distributed throughout the test. This suggests that, because the BIDR-IM scale is a self-report measure, it too would be subject to CMV resulting from socially desirable responding in the same manner, as self-report measures personality is prone to in high stakes selection contexts. The more readily the construct being assessed can be identified, from a reading of the items, the more likely it is for socially desirable responding to occur (McFarland et al., 2002). To deal with these concerns, rather than grouping them together, the 20 IM items could be randomly included in a 'bespoke' questionnaire containing distractor items. The distractor items selected should not be related to impression management. Rosse, Stecher, Miller, and Levin (1998) used a similar approach when using the IM scale of the BIDR in their research. This bespoke nature of the modified BIDR-IM scale, or similar measures, will arguably help to minimise a socially desirable responding effect from contaminating the accuracy of the IM measure used. The research reviewed suggests that the objective of minimising faking good in high stakes personality assessments can be achieved by making moral standards salient in the assessment procedure combined with the inherent self-awareness effect that comes from completing a self-report measure. In addition, the detection of job applicants who still fake good can be achieved by using an impression management measure that does not group the items in the measure together, as a procedural measure for dealing with the problem of item transparency.

## **Conclusion**

The extant research reviewed dealing with the use of personality measures in the assessment of job applicants, in high stakes contexts, clearly indicates that faking good on personality measures is a real problem when it comes to the applied context of high stakes employee selection. Resolving this issue is of major practical importance in such a context. According to Drasgow, Stark, Chernyshenko, Nye, Hulin, and White (2012, p. 2) "intentional distortion can severely undermine the utility of measures for personnel selection". The use of procedural controls in personality assessments, which make moral standards salient as well as bringing about a state of objective self-awareness, will act to prevent moral hypocrisy in a high stakes employee selection setting. Thus, it should be possible to minimise its occurrence, thereby leading to more accurate personality assessments of job candidates and better employee selection outcomes, particularly in high stakes contexts. However, even with procedural precautions, it will still be necessary to use an impression management measure, similar to the bespoke measure described earlier. This approach should help to ensure that those who persist in faking good, in spite of the procedural measures taken to eliminate moral hypocrisy, are detected.

## References

- Arthur, W., Glaze, R. M., Villado, A. J., and Taylor, J. E. (2010). The magnitude and extent of cheating and response distortion effects on unproctored internet based tests of Cognitive Ability and Personality. *International Journal of Selection and Assessment*, 18(1), 1-16.
- Bäckström, M., Björklund, F., and 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, 335-344.
- Bangerter, A., Roulin, N., and König, C. J. (2012). Personnel selection as a signaling game. *Journal of Applied Psychology*, 97(4), 719.
- Barrick, M. R., and Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44(1), 1-26.
- Barrick, M. R., and Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, 81(3), 261-272.
- Barrick, M. R., Mount, M. K., and 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(12), 9-30.
- Batson, C. D. (2008). Moral masquerades: Experimental exploration of the nature of moral motivation. *Phenomenology and the Cognitive Sciences*, 7(1), 51-66.
- Batson, C. D., Kobrynowicz, D., Dinnerstein, J. L., Kampf, H. C., and Wilson, A. D. (1997). In a very different voice: unmasking moral hypocrisy. *Journal of Personality and Social Psychology*, 72(6), 1335-1348.
- Batson, C. D., Thompson, E. R., Seufferling, G., Whitney, H., and Strongman, J. A. (1999). Moral hypocrisy: appearing moral to oneself without being so. *Journal of Personality and Social Psychology*, 77(3), 525-537.
- Bess, T. L. and Harvey, R. J.. (2002). Bimodal Score Distributions and the Myers–Briggs Type Indicator: Fact or Artifact? *Journal of Personality Assessment*. 78(1), 176–186.
- Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., and Smith, M. A. (2006). A Meta Analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment*, 14(4), 317-335.
- Block, J. (2010). The five-factor framing of personality and beyond: Some ruminations. *Psychological Inquiry*, 21(1), 2-25.

- Burns, G. N., and Christiansen, N. D. (2011). Methods of measuring faking behaviour. *Human Performance*, 24(4), 358-372.
- Campbell, D. T., and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81-105.
- Chan, D. (2009). So why ask me? Are self-report data really that bad. In Lance, Charles E.; Vandenberg, Robert J.(Eds.). *Statistical and Methodological Myths and Urban Legends: Doctrine, verity and fable in the organizational and social sciences*. New York: Routledge, pp. 309-336
- Collins, J. M., Schmitt, F. L., Sanchez-Ku, M., Thomas, L., McDaniel, M. A., and Le, H. (2003). Can basic individual differences shed light on the construct meaning of assessment center evaluations? *International Journal of Selection and Assessment*, 11, 17–29.
- Conn, S. and Rieke, M. L. (Eds.). (1994). The 16PF fifth edition technical manual. Champaign, IL: Institute for Personality and Ability Testing.
- Connelly, B. S., and Chang, L. (2016). A meta-analytic multitrait multirater separation of substance and style in social desirability scales. *Journal of Personality*, 84, 319-334.
- Costa Jr., Paul T. and McCrae, Robert R. (1992). *NEO Personality Inventory Revised (NEO PI-R)*. Florida: Psychological Assessment Resources Inc.
- Cronbach, L. J., and Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281.
- Crowne, D. P., and Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24(4), 349-357.
- DeYoung, C. G., Peterson, J. B., and Higgins, D. M. (2002). Higher-order factors of the Big Five predict conformity: Are there neuroses of health? *Personality and Individual Differences*, 33(4), 533-552.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41(1), 417-440.
- Dilchert, S., and Ones, D. S. (2012). Application of preventive strategies. In Zeigler, M., MacCann, C., and Roberts, R. (Eds.). *New Perspectives on Faking in Personality Assessment*. New York: Oxford Books, pp. 177-200.
- Dragow, F., Stark, S., Chernyshenko, O. S., Nye, C. D., Hulin, C. L., and White, L. A. (2012). *Development of the Tailored Adaptive Personality Assessment System (TAPAS) to support Army personnel selection and classification decisions*. Illinois: Urbana, Dragow Consulting Group.

- Duval, T. S., and Lalwani, N. (1999). Objective self-awareness and causal attributions for self-standard discrepancies: Changing self or changing standards of correctness. *Personality and Social Psychology Bulletin*, 25(10), 1220-1229.
- Ellingson, J. E. (2012). People fake only when they need to fake. In Ziegler, M., MacCann, C., and Roberts, R. (Eds.), *New Perspectives on Faking in Personality Assessment*, New York: Oxford University Press, pp. 19-33.
- Ellingson, J. E., Heggstad, E. D., and Makarius, E. E. (2012). Personality retesting for managing intentional distortion. *Journal of Personality and Social Psychology*, 102(5), 1063-1076.
- Ellingson, J. E., Smith, D. B., and Sackett, P. R. (2001). Investigating the influence of social desirability on personality factor structure. *Journal of Applied Psychology*, 86(1), 122-133.
- Embretson (Whitely), S. E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin*, 93(1), 179-197.
- Embretson, S. E. (2007). Construct validity: A universal validity system or just another test evaluation procedure? *Educational Researcher*, 36(8), 449-455.
- Eysenck, H. J. (1968). Eysenck personality inventory manual. *San Diego: Educational and Industrial Testing Service*.
- Fahey, G., (2017). *The Construct Validity of the NEO PI-R Personality Inventory in High Stakes Employee Selection*. (Doctoral Thesis). Retrieved from [http://doras.dcu.ie/21900/1/GerryFaheyFULL\\_THESIS\\_2.0.pdf](http://doras.dcu.ie/21900/1/GerryFaheyFULL_THESIS_2.0.pdf) (Accessed: 12 November 2018)
- Fan, J., Gao, D., Carroll, S. A., Lopez, F. J., Tian, T. S., and Meng, H. (2012). Testing the efficacy of a new procedure for reducing faking on personality tests within selection contexts. *Journal of Applied Psychology*, 97(4), 866-880.
- Fischbacher, U., and Föllmi-Heusi, F. (2013). Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association*, 11(3), 525-547.
- Griffith, R. L., and Converse, P. D. (2012). The rules of evidence and the prevalence of applicant faking. In Ziegler, M., MacCann, C., and Roberts, R. (Eds.), *New Perspectives on Faking in Personality Assessment*, New York: Oxford University Press, pp. 34-52.
- Griffith, R. L., Chmielowski, T., and Yoshita, Y. (2007). Do applicants fake? An examination of the frequency of applicant faking behaviour. *Personnel Review*, 36(3), 341-355.
- Hathaway, S. R., and McKinley, J. C. (1989). *MMPI-2: Minnesota Multiphasic Personality Inventory-2: manual for administration and scoring*. Minneapolis, Minn.: University of Minnesota Press.

- Highhouse, S. (1998). Understanding and improving job-finalist choice: The relevance of behavioral decision research. *Human Resource Management Review*, 7(4), 449-470.
- Highhouse, S. (2002). Assessing the candidate as a whole: A historical and critical analysis of individual psychological assessment for personnel decision making. *Personnel Psychology*, 55(2), 363-396.
- Hogan, J., Barrett, P., and Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology*, 92(5), 1270-1285.
- Hogan, J., Hogan, R., and Kaiser, R. B. (2010). Management derailment. *American Psychological Association Handbook of Industrial and Organizational Psychology*, 3, 555-575.
- Hogan, R. (2005). In defense of personality measurement: New wine for old whiners. *Human Performance*, 18(4), 331-341.
- Hogan, R., Hogan, J., and Roberts, B. W. (1996). Personality measurement and employment decisions: Questions and answers. *American psychologist*, 51(5), 469-477.
- Holden, R. R. (2008). Underestimating the effects of faking on the validity of self-report personality scales. *Personality and Individual Differences*, 44(1), 311-321.
- Hollenbeck, G. P. (2009). Executive selection—What's right... and what's wrong? *Industrial and Organizational Psychology*, 2(2), 130-143.
- Hough, L. M. (1998). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance*, 11(2-3), 209-244.
- Hough, L. M., and Oswald, F. L. (2008). Personality testing and industrial–organizational psychology: Reflections, progress, and prospects. *Industrial and Organizational Psychology*, 1(3), 272-290.
- Institute of Psychometric Coaching (2017). Retrieved from <http://www.psychometricinstitute.com.au/Personal-Psychometric-Coaching.html> (Accessed: 21 January 2017).
- Jeong Y., Christiansen N., Robie C., Kung M-C., and Kinney T. (2017) Comparing applicants and incumbents: Effects of response distortion on mean scores and validity of personality measures. *International Journal of Selection and Assessment*, 25, 31-315.
- Judge TA, and Kammeyer-Mueller JD. (2007). Personality and career success. In Peiperl, L., and Gunz, H. (Eds.), *Handbook of career studies*. Thousand Oaks, CA: Sage.

Judge, T. A., Higgins, C. A., Thoresen, C. J., and Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, 52(3), 621-652.

Judge, T. A., Bono, J. E., Ilies, R., and Gerhardt, M. W. (2002). Personality and leadership: a qualitative and quantitative review. *Journal of Applied Psychology*, 87(4), 765-780.

Kiefer, C. and Benit, N. (2016). What is Applicant faking behavior? A review on the current state of theory and modeling techniques. *Journal of European Psychology Students*, 7(1), 9–19.

Komar, S., Brown, D. G., Komar, J. A., and Robie, C. (2008). Faking and the validity of conscientiousness: A Monte Carlo investigation. *Journal of Applied Psychology*, 93, 140–154.

Kurtz, J. E., Tarquini, S. J., and Iobst, E. A. (2008). Socially desirable responding in personality assessment: Still more substance than style. *Personality and Individual Differences*, 45(1), 22-27

Lance, C. E., Dawson, B., Birkelbach, D., and Hoffman, B. J. (2010). Method effects, measurement error, and substantive conclusions. *Organizational Research Methods*, 13(3), 435-455.

Landers, R. N., Sackett, P. R., and Tuzinski, K. A. (2011). Retesting after initial failure, coaching rumors, and warnings against faking in online personality measures for selection. *Journal of Applied Psychology*, 96(1), 202 - 210

LaPiere, R. T. (1934). Attitudes vs. actions. *Social Forces*, 13(2), 230-237.

Li, A., and Bagger, J. (2006). Using the BIDR to distinguish the effects of impression management and self-deception on the criterion validity of personality measures: A meta-analysis. *International Journal of Selection and Assessment*, 14(2), 131-141.

Lönnqvist, J. E., Irlenbusch, B., and Walkowitz, G. (2014). Moral hypocrisy: impression management or self-deception? *Journal of Experimental Social Psychology*, 55, 53-62.

MacCann, C., Ziegler, M., and Roberts, R. D. (2012). Faking in personality assessment – reflections and recommendations. *New Perspectives on Faking in Personality Assessment*, New York: Oxford University Press, pp. 309-329.

Mayer, J. D., Salovey, P., and Caruso, D. R. (2008). Emotional intelligence: New ability or eclectic traits? *American Psychologist*, 63, 503–517.

Mazar, N., Amir, O., and Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, 45(6), 633-644.

- McCrae, R. R. (2010). The place of the FFM in personality psychology. *Psychological Inquiry*, 21(1), 57-64.
- McCrae, R. R., and Terracciano, A. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology*, 88(3), 547-561.
- McFarland, L. A., and Ryan, A. M. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology*, 85, 812-821.
- McFarland, L. A., Ryan, A. M., and Ellis, A. (2002). Item placement on a personality measure: Effects on faking behaviour and test measurement properties. *Journal of Personality Assessment*, 78(2), 348-369.
- Messick, S. (1995). Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., and Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, 60(3), 683-729.
- Morin, A. (2011). Self-awareness part 1: Definition, measures, effects, functions, and antecedents. *Social and Personality Psychology Compass*, 5(10), 807-823.
- Mount, M. K., Barrick, M. R., and Stewart, G. L. (1998). Five-factor model of personality and performance in jobs involving interpersonal interactions. *Human performance*, 11(2-3), 145-165
- Niessen, S.M., Meijer, R.R., and Tendeiro, J. N. (2017). Measuring non-cognitive predictors in high-stakes contexts: The effect of self-presentation on self-report instruments used in admission to higher education. *Personality and Individual Differences*, 106, 183-189.
- Oh, I. S., Kim, S., and Van Iddekinge, C. H. (2015). Taking it to another level: Do personality-based human capital resources matter to firm performance? *Journal of Applied Psychology*, 100(3), 935-947
- Ones, D. S., Viswesvaran, C., and Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology*, 81(6), 660-679.
- Ones, D. S., Dilchert, S., Viswesvaran, C., and Judge, T. A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology*, 60(4), 995-1027.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46(3), 598-609.

- Paulhus, D. L. (1998). Paulhus deception scales (PDS): *The balanced inventory of desirable responding* – 7. North Tonawanda, NY: Multi-Health Systems.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. *The Role of Constructs in Psychological and Educational Measurement*. Abingdon, Oxfordshire: Routledge, pp. 49-69.
- Perinelli, E., and Gremigni, P. (2016). Use of social desirability scales in clinical psychology: a systematic review. *Journal of clinical psychology*, 72(6), 534-551.
- Ployhart, R. E., Schmitt, N., and Tippins, N. T. (2017). Solving the Supreme Problem: 100 Years of Selection and Recruitment at the Journal of Applied Psychology. *Journal of Applied Psychology*, 102(3), 1-14.
- Podsakoff, P. M., MacKenzie, S. B., and Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, 63, 539-569.
- Pruckner, G. J., and Sausgruber, R. (2013). Honesty on the streets: A field study on newspaper purchasing. *Journal of the European Economic Association*, 11(3), 661-679.
- Pryor, J. B., Gibbons, F. X., Wicklund, R. A., Fazio, R. H., and Hood, R. (1977). Self-focused attention and self-report validity. *Journal of Personality*, 45(4), 513-527.
- Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., and Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, 2(4), 313-345.
- Rosse, J. G., Stecher, M. D., Miller, J. L., and Levin, R. A. (1998). The impact of response distortion on pre-employment personality testing and hiring decisions. *Journal of Applied Psychology*, 83(4), 634-644.
- Roulin, N., and Bourdage, J. S. (2017). Once an impression manager, always an impression manager? Antecedents of honest and deceptive Impression Management use and variability across multiple job interviews. *Frontiers in Psychology*, 8, 1-13.
- Sackett, P. R.. (2012). Faking in personality assessment – where do we stand? In Ziegler, M., MacCann, C., and Roberts, R. D. *New Perspectives on Faking in Personality Assessment*. New York: Oxford University Press, pp. 330-344.
- Sackett, P. R., Schmitt, N., Ellington, J. E., and Kabin, M. B. (2001). High-stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative-action world. *American Psychologist*, 56, 302-318.

- Salgado, J. F. (2003). Predicting job performance using FFM and non FFM personality measures. *Journal of Occupational and Organizational Psychology*, 76(3), 323-346.
- Salgado, J. F. (2005). Personality and social desirability in organizational settings: practical implications for work and organizational psychology. *Papeles del Psicólogo*, 26(S 124), 115-128.
- [https://www.researchgate.net/publication/242537719\\_Personality\\_and\\_social\\_desirability\\_in\\_organizational\\_settings\\_Practical\\_implications\\_for\\_work\\_and\\_organizational\\_psychology](https://www.researchgate.net/publication/242537719_Personality_and_social_desirability_in_organizational_settings_Practical_implications_for_work_and_organizational_psychology)  
(Accessed: 12 November 2018)
- Salgado, J. F. (2016). A Theoretical Model of Psychometric Effects of Faking on Assessment Procedures: Empirical findings and implications for personality at work. *International Journal of Selection and Assessment*, 24(3), 209-228.
- Schmit, M. J., and Ryan, A. M. (1993). The Big Five in personnel selection: Factor structure in applicant and non-applicant populations. *Journal of Applied Psychology*, 78(6), 966-974.
- Shaver, P. R., Brennan, K. A., Robinson, J. P., Shaver, P. R., and Wrightsman, L. S. (1991). Measures of personality and social psychological attitudes. Robinson, J.P., Shaver, P.R., and Wrightsman L.S.,(Eds.). *Measures of Depression and Loneliness*, 1, San Diego, California: Academic Press, pp. 212-215.
- Shu, L. L., Gino, F., and Bazerman, M. H. (2011). Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated forgetting. *Personality and Social Psychology Bulletin*, 37(3), 330-349.
- Shu, L. L., Mazar, N., Gino, F., Ariely, D., and Bazerman, M. H. (2012). Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *Proceedings of the National Academy of Sciences*, 109(38), 15197-15200.
- Silvia, P. J., and Duval, T. S. (2001). Objective self-awareness theory: Recent progress and enduring problems. *Personality and Social Psychology Review*, 5, 230-241.
- Smith, D. B., and Ellingson, J. E. (2002). Substance versus style: A new look at social desirability in motivating contexts. *Journal of Applied Psychology*, 87(2), 211-219.
- Sollman, M. J., and Berry, D. T. (2011). Detection of inadequate effort on neuropsychological testing: A meta-analytic update and extension. *Archives of Clinical Neuropsychology*, 26, 774-789.
- Wise, S. L., and DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment* 10 (1), 1-17.

Ziegler, M., MacCann, C., and Roberts, R. D. (2012). Faking: Knowns, unknowns, and points of contention. In Ziegler, M., MacCann, C., and Roberts, R.D. (Eds.). *New Perspectives on Faking in Personality Assessment*. New York: Oxford University Press, pp. 3-16.