



Testing on the Internet: Faking in a Web-Based Self-Administered Personality Measure

Udo Konradt¹, Sebastian Sypersek¹, Guido Hertel²

¹University of Kiel
²University of Muenster

ABSTRACT

This experimental study examined the effect of faking in a self-administered occupational personality questionnaire on the World Wide Web and explored whether the degree of faking is related to Self-monitoring and login times. Using a between subjects design, employees were instructed to "answer honestly", to "fake good", or to answer as if applying for a job. Results revealed faking in the web-based personality questionnaire for two out of five dimensions (i. e., conscientiousness and self-motivation). Age was found to moderate the degree of faking with older people showing a higher degree of faking. Controlling for age and level of education, results revealed that high self-monitors demonstrated more faking compared to low self-monitors. Finally, slow responders showed higher degrees of faking compared to fast responders. We also discuss implications for online personnel selection and suggestions for future research.

Keywords: Computerized Testing, internet-based testing, faking, personnel selection, self-monitoring.

1 Introduction

From the beginning of personality research, evidence has been found that personality measures are susceptible to socially desirable responding (Braun, 1963; Meehl & Hathaway, 1946). Faking or intentional response distortion refers to an attempt to manipulate responses to a personality measure in a job application process. Ample evidence exists that people who are instructed to "fake good" show elevated scores compared to conditions when they are not instructed to do so (Birkeland, Manson, Kisamore, Brannick & Smith, 2006; Hogan, Barrett, & Hogan, 2007; Hough & Furnham, 2003; Smith & Robie, 2004).

Faking in personality measurement has various consequences regarding the validity of personality measures and has substantially initiated research (e. g., Edens & Arthur, 2000; Ellingson, Sackett, & Hough, 1999; Ones, Viswesvaran, & Reiss, 1996; Richman, Kiesler, Weisband, & Drasgow, 1999; Viswesvaran & Ones, 1999; Zickar & Drasgow, 1996; Zickar & Robie, 1999). A main concern is the potential deterioration of the criterion-related validity of personality measures (Ellingson et al., 1999). Consequently, scholars are interested in a better understanding of the antecedents which prevent faking and therefore evoke veridical self-reports (Anderson, Silvester, Cunningham-Snell, & Haddleton, 1999; McFarland & Ryan, 2000; Richman et al., 1999; Snell, Sydell, & Lueke, 1999).

A strong and ongoing trend in human resource (HR) management processes and systems is to use computers for

applicant recruitment and personnel selection (Barak & English, 2002; Hertel, Konradt, & Orlikowski, 2003; Konradt & Hertel, 2004; Naglieri, Drasgow, Schmit, Handler, Prifitera, Margolis & Velasquez, 2004). Despite the increase in computer-based personality measures and various possibilities of test applications via the Internet, there is a paucity of research on these issues (Cardy & Miller, 2003; Dwight & Feigelson, 2000; Tippins, Beaty, Drasgow, Gibson, Pearlman, Segall & Shepherd, 2006). Specifically, the question was raised whether the degree of faking might be influenced by the administration mode, e. g., if faking occurs to a similar degree in computerized and paper-pencil measures (Dwight & Feigelson, 2000). The cardinal problem with web-based testing is the lack of control over the setting during the test period which might deteriorate the test objectivity and might enhance the applicant's motivation to fake. Evidence was found that, in a computer-assisted self-administration test, the impression management scores were more positively biased in a distant condition compared to a proximate condition (Moon, 1998; Richman et al., 1999). These results suggest that remote-access tests may lead to more "biased" self-judgments. On the other hand, computerized testing offers the benefit of using latencies to detect faking behavior (Holden, 1998). Despite the worldwide use of computers in personnel selection, very little research on HR management and related fields has focused on how these conditions change the answering behavior of applicants (Cardy & Miller, 2003; Dwight & Feigelson, 2000; Konradt, Lehmann, Böhm-Rupprecht, & Hertel, 2003). The aim of this study was to extend the knowledge of the subject and to provide a foundation for practice-related recommenda-

tions. Specifically, this study explores faking in a self-administered web-based personality measure in relation to both person-related factors (i. e., Self-monitoring) and response latencies as a faking indicator.

2 Empirical Background and Hypotheses

2.1 Faking in Computerized Personality Testing

In personnel selection, faking is based on the motivation of a respondent to match the expected self-presentation to a job (Richman et al., 1999). Socially desired but untrue answers are related to impression management, which refers to a process in which individuals attempt to influence the opinions or perceptions of an external audience (Rosenfeld, Giacalone, & Riordan, 1995). Examining the construct of social desirability, Paulhus (1984) distinguishes between the components *impression management* (the deception of others) and *self-deception*. Evidence for the occurrence of impression management techniques has been found in assessment interviews (Gilmore & Ferris, 1989; Stevens & Kristof, 1995), performance appraisals (Wayne & Ferris, 1990; Wayne & Liden, 1995), and leadership behavior (Wayne & Green, 1993). Specifically, faking has been studied in computerized offline testing settings in which tests are performed under controlled conditions. In a meta-analysis on the effect of computerized testing on response bias, Dwight and Feigelson (2000) found a significant though small effect. However, despite the growing interest in web-based personnel recruitment and assessment (Barak & English, 2002; Bartram, 1999), no study has addressed faking using online self-administered instruments. Since the majority of studies on computerized testing in offline settings found evidence for faking, it was hypothesized:

Participants who are instructed to fake in a self-administered online personality measure show higher scores compared to participants who were not instructed to fake (*hypothesis 1*).

Furthermore, it is of interest if particular personality scales are easier to fake than others. Drawing on paper-pencil tests, Viswesvaran and Ones (1999) showed that all of the Big Five factors could equally be faked. Contrarily, McFarland and Ryan (2000) demonstrated that personality factors can be faked to different extents. Like other studies using the Big Five (e. g., Paulhus, Bruce, & Trapnell, 1995; Topping & O’Gorman, 1997), they provided evidence that, compared to the other scales, participants faked the Openness scale to a lower and Extraversion scale to a greater extent. Moreover, Dalen, Stanton and Roberts (2001) varied the job profile information given to candidates who were asked to fake a personality questionnaire. They demonstrated that candidates faked on dimensions which were relevant to the job profile, indicating a differential effect of faking. Because of these conflicting results, we put forth an open *Research Question*:

Can different personality scales of a web-based personality questionnaire equally be faked?

2.2 Antecedents of Faking Behavior

Among the antecedents of faking behavior, Self-monitoring has also been considered to be crucial (Day, Unckless, Schleicher, & Hiller, 2002). Referring to the control of self-presentational behavior (Snyder, 1974), the Self-monitoring concept has been extensively empirically examined (see Gangestad & Snyder, 2000, for a review). High self-monitors attend and are influenced by situational cues which guide their self-presentation, while low self-monitors display behavior consistent with their own feelings. Therefore, high self-monitors control their behavior and adapt to a situation more accurately (Snyder, 1987). Self-monitoring is also positively related to social desirability and impression management (Leary & Kowalski, 1990; Paulhus, 1984; Rosenfeld et al., 1995). While both constructs are theoretically distinct (see Miller & Cardy, 2000), they are empirically related. For example, high self-monitors can anticipate what their counterparts want and look for (Turnley & Bolino, 2001). McFarland and Ryan (2000) demonstrated that high self-monitors instructed to “fake good”, i. e. to give a positive impression, increase their Extraversion scores to a greater extent than low self-monitors. Thus, consistent with the effect of Self-monitoring on social behavior, it is hypothesized:

High self-monitors can enhance their positive impression (fake good) on the web better than low self-monitors (*hypothesis 2*).

2.3 Detecting Fakers

Many faking studies explored ways to detect tendencies to answer in a socially desired way (Austin, 1992; Kelly & Greene, 1989; Zickar & Drasgow, 1996) and several strategies have been suggested for coping with this problem (Christiansen, Goffin, Johnston, & Rothstein, 1994; Lanyon, 1993). In computer-administered questionnaires, technical indicators of faking behavior were primarily investigated. In this process, login time was used as an indicator of cognitive answering processes. For example, complex, difficult or invasive items should lead to longer latencies. Consequently, Holden (1998) reported that faking leads to slower responses. In accordance with Holden (1998), it was expected:

Applicants faking good (i. e., faking to put themselves in a positive light) will take longer response times than honest job candidates (*hypothesis 3*).

3 Method

3.1 Sample

Four-hundred-sixty-eight German employees between 18 and 40 years of age were contacted to participate in the study. The response rate was 74% (354 data sets). From

these data sets, 42 were excluded from analysis due to extensive missing data¹. 155 females and 157 males remained with an average age of 29.6 years ($SD = 5.9$). 44 participants (14.1%) had primary school education ("Hauptschulabschluss"), 102 (32.7%) had completed secondary school ("Mittlerer Bildungsabschluss"), 78 (28.2%) had a higher secondary school education ("Hochschulreife"), and 88 (28.2%) had attended university. Participants were randomly assigned to one of three experimental conditions (see procedure section). Groups did not differ significantly with regard to gender, age, level of education, and occupational status ($F_s < 1, n. s.$). No participant was employed in the work area (i. e., call center agent) which was used for the setting.

3.2 Measures

Personality Tests. We explored the faking of online assessments using the Call Center Aptitude Test (Konradt, Hertel, & Joder, 2003, 2004), which is a computer-based test battery for the selection of call center agents. This test battery was used because it was developed as a web-based assessment tool for personnel selection and has good psychometric properties. The personality test module included twenty items addressing five personality traits: the willingness to learn (three items, e. g., 'I am fascinated by complex subjects'), social competence (four items, e. g., 'It is rather easy for me to reveal my opinion in an unknown situation'), self-motivation (three items, e. g., 'If I don't succeed, I will increase my efforts'), stress resistance (five items, e. g., 'When I have to work to a deadline under pressure, I perform less than usual' (reversed)), and conscientiousness (five items, e. g., 'Responsible tasks are important to me'). Items were answered on five-point scales (from 'not at all' to 'completely') with varying polarity. After recoding the reversed items, scale reliabilities (alpha) were between .66 and .75 (see Table 2)².

Self-monitoring. Self-monitoring was assessed by a German adaptation of Snyder's (1974) 25-item version of the Self-Monitoring Scale (Mielke & Kilian, 1990). Three items with the highest selectivity score were selected from each of the three subscales ('social abilities', 'inconsistency', and 'social comparing'). The resulting nine items were

¹ Cut-off scores were used as a plausibility check. However, the main results were basically the same even when all of the participants were included who returned the questionnaire.

² To assess the dimensionality of the personality test, a confirmatory factor analysis was performed on the data merged from the three experimental conditions. Analyses were conducted with structural equation modeling techniques employing the maximum likelihood procedure, and using the AMOS 5.01 software package (Arbuckle, 2003). A check for the distributional assumptions of maximum likelihood estimation revealed that univariate and multivariate skewness and kurtosis were within the expected range (cf. Tabachnik & Fidell, 2006). Evidence that the five-factor model acceptably fits the data is provided by the root-mean square error of approximation (RMSEA) of .07 (cf. Hu & Bentler, 1999). Similarly, the goodness of fit index (GFI = 0.87, adjusted GFI = .083) and the normed fit index (NFI = 0.77) indicated an acceptable model fit. Employing chi-square difference tests demonstrated that this model provided a fit superior to that of the one-factor model ($\Delta X_{(17)} = 2096.30, p < .001$).

scored on a Likert-type scale ranging from 1 ('strongly disagree') to 7 ('strongly agree'). A confirmatory factor analysis of the scale generally supported the three-factor model (GFI = 0.89, adjusted GFI = 0.80, NFI = 0.77, TLI = 0.79, RMSEA = 0.14). Cronbach's alpha was .82.

Controls. To conduct conservative tests of the hypotheses, we controlled for age (in years) and education level: (1) completed Primary school, (2) completed Secondary school, (3) completed High school, and (4) obtained a university degree.

3.3 Procedure

Three-hundred-fifty-four employed adults who were part of an extensive commercial panel for online surveys were selected randomly. They received six Euros for their participation in the study. Participants were invited to visit a web site and asked to complete the study within one week. The web site provided (a) information about the job (title, description and the qualification requirements which are relevant for the job in terms of the five personality dimensions), (b) the instruction either to answer honestly (condition 1), to fake good (condition 2), or to answer like an applicant (condition 3) and (c) the web-based questionnaire³. As commonly used in faking studies (see Hough & Furnham, 2003; Smith & Robie, 2004), the instruction of the fake good condition was to maximize chances of employment. In the application condition, participants were told to answer like an applicant; and in the honest condition, they were asked to answer as honestly as possible. Participants were instructed to engage only in this task while logged on. Response latency in answering, i. e., how long they spent completing the questionnaire (login time) was automatically recorded. The average login time was 13.3 min ($SD = 4.6$ min) ranging between 7 and 30 minutes.

4 Results

Table 1 presents the means, standard deviations and reliabilities of the personality scales for each condition. The reliability estimates ranged from .66 to .75 and mostly met the level of .70 required for the internal consistency of personality measures (cf. Nunnally & Bernstein, 1994)⁴. The overall reliability of the personality test battery was $r = .90$. As shown in Table 2, all of the five personality measures showed significant correlations ($p < .01$), but were unrelated to other individual variables. The only significant relation was age. Correlations were also calculated using the experimental group variable as a covariate. However, the zero-order and first-order correla-

³ As the instructions were very easy to understand and had been used in several studies before (see Hough & Furnham, 2003), we excluded a manipulation check.

⁴ Internal reliabilities below the limit of .70 seem to be mainly due to the small scale sizes. Following the Spearman Brown prophecy formula, an extension on eight items of each scale would lead to reliabilities (r' tt) between .76 and .89.

tions were very similar, hence only zero order correlations are presented.

To test Hypothesis 1, which predicted that people are able to fake a self-administered online personality measure, we ran an ANOVA on the mean scores of the overall personality questionnaire score, with condition as the factor.⁵ Results yielded a significant main effect ($F(2, 309) = 3.00, p < .05$), explaining only two percent ($\eta^2 = 0.02$) of the total variance and providing support for Hypothesis 1. Post hoc t-tests showed significant differences in Self-motivation ($t(2) = 4.74, p < .01$) and Conscientiousness, $t(2) = 3.61, p < .05$ between the conditions. Here, the means in the application condition and the fake condition were nearly the same (for Self-motivation: $M = 5.52, SD = 0.91$ vs. $M = 5.53, SD = 1.10$, and for Conscientiousness: $M = 5.96, SD = 0.70$ vs. $M = 5.93, SD = 0.83$), but significantly higher ($p < .05$) than in the honest condition (for self-motivation: $M = 5.15, SD = 1.06$, and for Conscientiousness: $M = 5.70, SD = 0.81$). The effect sizes between the honest and the application condition were $d = 0.36$ for Self-motivation and $d = 0.35$ for Conscientiousness, indicating a low to medium effect (cf. Cohen, 1988). Thus, Hypothesis 1 was supported.

Hypothesis 2 proposed that high self-monitors are more likely to fake compared to low self-monitors. We used moderated hierarchical regression analyses to investigate the Self-monitoring effects on faking separately for the three conditions. In the first step, age and level of education were entered as control variables. Secondly, the three Self-monitoring subscale scores were entered simultaneously. As shown in Table 3, age was statistically significant for both the faking and application condition, with positive betas greater than .33 and a significant R^2 of .11 (adjusted $R^2 = .10$), and .12 (adjusted $R^2 = .11$), respectively. Under the honest condition, age had no effect ($\beta = .06, n. s.$). Regarding main effects, the betas of the two Self-monitoring scales 'social abilities' and 'inconsistency' were significantly related to the mean score on personality scales, in all three conditions, with positive betas for social abilities (betas ranging between .15 and .34) and significantly negative betas for inconsistency (betas between $-.12$ and $-.16$). Social comparison processes had no effect. The increment of the Self-monitoring variables with regard to the explained variance was statistically significant in the honest ($\Delta R^2 = .28, p < .001$) and the application condition ($\Delta R^2 = .14, p < .001$), but not in the fake condition ($\Delta R^2 = .05, n. s.$), showing partial support for Hypothesis 2.

Hypothesis 3 stated that applicants faking good have longer response times than honest job candidates. To explore the relation between faking and response latencies, participants who were asked to fake were compared to those who were asked to answer honestly. A median split across all of the conditions on the login time created

a fast responder group ($M = 10.1, SD = 1.5$) which differed significantly from a slow responder group ($M = 17.2, SD = 4.0, t(311) = 44.92, p < .001$). The mean group difference was two minutes. To test Hypothesis 3, an ANOVA⁶ was conducted with response latency as the dependent variable. A main effect was found ($F(2, 310) = 5.12, p < .01, \eta^2 = 0.37$). Post hoc t-tests showed significantly longer login times in the fake condition compared to the honest condition ($t(208) = 3.13, p < .01$), thereby supporting Hypothesis 3. As older people responded more slowly, we ran an ANCOVA on the login times, with condition as a factor and age as a covariate. Again Hypothesis 3 was supported by yielding a significant main effect ($F(2, 310) = 4.93, p < .01, \eta^2 = 0.32$), with longer login time in the fake compared to the honest condition when controlling for age, $t(208) = 3.12, p < .01$.

Since the login time differed between the fake and the honest conditions, it was examined if login time can be used to discriminate between honest and dishonest participants. Discriminant analyses were conducted with the condition (faker vs. non-faker) as the dependent variable and login time as a predictor, revealing a significant result ($\chi^2(1) = 9.52, p < .01$). Thus, 57.1% of the participants were correctly classified according to their original group by the discriminant function⁶. As regards to the relative effect of login time on the faking of participants, the standardized canonical discriminant function coefficient was .22, reflecting a small function of the login time as a faking predictor. Given the significant though small effects, these findings provide partial support that login time can predict faking in an online setting.

5 Discussion

The aim of this study was to determine whether individuals are able to fake self-administered Internet personality questionnaires and whether Self-monitoring and login time can predict the level of faking. Results indicated that two out of five personality scales (i. e., Self-motivation and Conscientiousness) showed higher mean scores as a result of faking good. The mean scores of the other three scales (i. e., Willingness to learn, Social competence and Stress resistance) were higher after faking and the instruction given to the subject, although this effect was not significant. The differential inflation effect of the scales found in this study is in line with other studies (e. g., Dalen et al., 2001; Furnham, 1990) and suggests differences in the strategy of answering to the single dimensions. Participants seem not use a unilateral elevation strategy of consistently choosing the highest values on each item. Birkeland et al. (2006) meta-analyzed studies on faking responses in personality tests and demonstrated that applicants tend to be most concerned with inflating

⁵ Considering age as a covariate revealed a similar main effect for login time ($F(2,311) = 4.93, p < .01, \eta^2 = .03$).

⁶ A leave-one-out cross-validation procedure (Browne & Cudeck, 1989) was used to check the stability of the classification. In this procedure each subject is classified into one of the two groups according to the discriminant function computed from all the data, except the subject being classified. The proportion of misclassified subjects after removing the effect of each subject, one at a time, resulted in no shrinkage.

their scores on the conscientiousness and emotional stability dimensions. They conclude that responders "might view these constructs as being particularly desirable by employers and, thus, focus their attention on inflating their scores on these dimensions" (p. 327).

As an alternative explanation for the differential elevating effect, Sackett and Lievens (2008) suggested that instructed faking studies vary in terms of focusing on a specific job, on the workplace in general, or on a non-specified context (see also Birkeland et al., 2006). In our study, we focused on a specific job which might have allowed a differential elevation effect. The results of this study also show that the average effect of the faking manipulation was .30 standard deviations. In accordance with other studies on offline computer-administered tests within controlled environments (e. g., Birkeland et al. 2006; Richman et al., 1999; Viswesvaran & Ones, 1999), it can thus be concluded that the faking distortion in self-administered web-based questionnaires is rather small. Evidence suggests that people who are able to fake well do not necessarily do so in reality (Abrahams, Neumann, & Githens, 1971; Birkeland et al., 2006; Edens & Arthur, 2000; Hough & Schneider, 1996). Similarly, Viswesvaran and Ones (1999) showed an overestimation of naturally occurring levels of faking in personnel selection settings.

The fact that this study provides no evidence for additional faking effects in web-based compared to paper-and-pencil tests raises another important issue. Evans, Garcia, Garcia and Baron (2003) demonstrated that participants actually gave more veridical answers when the experimenter was removed from a face-to-face test setting. Moon (1998) investigated whether social desirability effects in computer-assisted self-administration tests are mitigated by simply varying the topological distance (i. e., between different cities). She found that impression management scores were significantly higher in a distant condition than in a more proximate condition (see Richman et al., 1999, for similar results). All in all, results suggest that negative presence effects might be compensated by a private and remote-access assessment test and may lead to more "unbiased" judgments. This argument is supported by computer-mediated communication research, which demonstrated that computer-mediated communication is not deteriorated by unconscious behavior. Rather, it is more prone to private self-awareness because self-related issues are usually more salient compared to issues related to others (Joinson, 2001; Matheson & Zanna, 1988). Nevertheless, the question whether respondents give more truthful answers in a computerized setting warrants further research.

In congruence with results from face-to-face settings (e. g., McFarland & Ryan, 2000; Turnley & Bolino, 2001), in a web-based testing high self-monitors present themselves more desirably than low self-monitors. Contrary to our prediction, Self-monitoring only adds a significant degree of variance to the test score in the honest and the application condition. In the fake condition, age and level of

education explained the main amount in criteria variance. This could be due to the type of the instrument which was used in this study. We assume that due to the high item transparency, participants could easily identify the "ideal profile", leading to low variances. Additionally, results seem to indicate that older people faked to a greater extent. An explanation is that age is positively related to job experience and job familiarity (Hough, Oswald, & Ployhard, 2001). Future research may thus consider the role of dimensional factors, such as item transparency (e. g., Kleinmann, Kuptsch, & Köller, 1996).

Although the effect strength of the relationship between login time and faking was small, our data emphasize the presence of a positive relationship. Hence, this result does not fully rule out Holden's (1998) notion that response latency can serve as an indicator for the resulting supplementary cognitive answering processes during faking. Regardless of the relationship revealed in our results, we suggest that additional individual and environmental differences may account for the time taken to answer. In their review on personality variables in work settings, Hough and Furnham (2003) concluded that response latency is correlated with personality variables, such as extraversion, neuroticism, and psychoticism. Moreover, they point out that response latency is correlated with job familiarity (see also Smith & Robie, 2004). A further complication is that test takers might differ in their reading ability or in their fluency of the test language, thus leading to longer test times. Taken together, this urges caution using login time as a measure of intentional distortion and as an indicator of faking.

5.1 Limitations

As with any study, this study has a number of limitations. First, a between-subjects design is used, resulting in relatively small effect sizes compared to studies using within-subjects designs, in which participants are given the chance to use a web-test several times (Edens & Arthur, 2000; Viswesvaran & Ones, 1999). Thus, the effect size and strength of faking may have been underestimated. Second, as the instructions were very easy to understand and had been used in several studies before (e. g., Hough & Furnham, 2003), we excluded a manipulation check. Former results clearly confirm faking in offline testing settings (Birkeland et al., 2006; Hough & Furnham, 2003; Smith & Robie, 2004), indicating that manipulation was successful. However, we could not fully rule out that the experimental groups do not act as they were told. Third, as a possible limitation, we have not yet examined variables which mediate faking behavior. In order to examine mediating variables, the motivation to participate (Arvey & Strickland, 1991) and the expected negative consequences of faking should be considered. Meta-analytic results yielding lower score deterioration in real job applications may be explained by a more meaningful context which corresponds with lower tendencies to fake (Edens & Arthur, 2000). Yet there have been only few studies on the impact of faking in computerized settings using job

applicants (see Van de Vijver & Harsveld, 1994, for an exception). In our study, methods or tactics to adapt the application situation were used which might be seen as a potential limitation. Finally, the results of our study indicate that login times could be used to identify fakers in web-based personality assessment. Like in web-based assessment settings, we were not able to measure average response latency on the item level, though this measure would be much more sensitive than login time. Future research should thus continue to investigate time to answer as an indicator of faking and to shed light on possible relations between response latencies on the item, construct, and test level.

5.2 Implications

There are several research and managerial implications in respect to the results of this study. There is still skepticism about whether the invisible net-applicant is "real" and produces valid and unbiased data (Barak & English, 2002; Bartram, 1999; Konradt & Hertel, 2004; Lievens & Harris, 2003; Tippins et al., 2006). The practical implications of this study are, first, that faking can occur under web-based testing but, second, that this faking effect is probably not higher than the effects reported for computer-based and paper-pencil measures. Since faking is demonstrated especially for high self-monitors, one way to avoid intentional distortion in personnel selection is to use indicators of objective behavior, such as login time. However, the results of this study show that login time may indicate faking only to some extent. Thus, a Self-monitoring together with a Social desirability scale or Lying scales can be used in order to identify dishonest respondents and to reduce response distortion. Building on the suggestions of Zickar and Drasgow (1996) and Ellingson, Sackett, and Hough (1999) of the limited success of these approaches, Schmitt and Oswald (2006) proposed removing applicants with high scores on faking indicators. Another important consideration is the significance of response distortion effects on procedural fairness. In many operational settings, web-based selection tools are used with a relatively low fixed cutoff score as part of initial screening and pre-selection. Sackett and Lievens (2008) argue that in such a setting, faking may erroneously cause a candidate to move to the next application stage, but that does not inhibit other candidates to respond honestly. Besides corrective response distortion reduction techniques, it seems important to direct attention to the reduction and prevention of response distortion. In this respect, the techniques to employ are, e. g., to warn candidates that faker will be identified and penalized, to request that candidates elaborate on their answers, and to use forced response formats (Sackett & Lievens, 2008). Finally, faking might also be regarded as an ability that reflects predictive validity. For example, Pauls and Crost (2005) demonstrated that general intelligence, self-reported efficacy and positive self-presentation, which are good predictors of job performance, were positively related to the amount of faking. Future research should thus pay more attention to the

interplay between preventive techniques for web-based testing (see Bank, 2003, for suggestions) and corrective response distortion techniques (i. e., the correction of values and removing candidates) in order to create conditions resulting in more valid responsiveness and predictive validity functions.

Finally, studies which have shown that impression management and faking behavior is moderated by various motivational factors (Nettlek & Leary, 2002) and situational characteristics (Robie, Born, & Schmit, 2001) indicate that faking behavior is best explained by a variety of skill dimensions (see also Hogan et al., 2007). Hence, in their multi-factorial model of faking, McFarland and Ryan (2000) argue that beliefs toward faking affects test scores and that attitudes regarding computers (Bratton & Newsted, 1995; Mahar, Henderson, & Deane, 1997) should be considered. As the mechanisms of faking under web-based testing are poorly understood, future research should continue to investigate the response distortion of candidates as a part of cognitive and motivational theories of organizational behavior.

6 References

- Abrahams, N. M., Neumann, I., & Githens, W. H. (1971). Faking vocational interests: Simulated versus real life motivation. *Personnel Psychology, 24*, 5–12.
- Anderson, N., Silvester, J., Cunningham-Snell, N., & Haddleton, E. (1999). Relationships between candidate self-monitoring, perceived personality, and selection interview outcomes. *Human Relations, 52*, 1115–1131.
- Arbuckle, J. (2003). *Amos 5.0 update to the Amos user's guide*. Chicago: SmallWaters.
- Arvey, R. D., Strickland, W., Drauden, G., & Martin, C. (1990). Motivational components of test-taking. *Personnel Psychology, 43*, 695–716.
- Austin, J. S. (1992). The detection of fake good and fake bad on the MMPI-2. *Educational and Psychological Measurement, 53*, 669–674.
- Bank, J. (2003). Erfahrungen und Positionen aus der Sicht eines psychometrischen Internet Serviceanbieters in den USA. In U. Konradt & W. Sarges (Eds.). *E-Recruitment und E-Assessment* (pp. 240–252). Göttingen: Verlag für Angewandte Psychologie.
- Barak, A., & English, N. (2002). Prospects and limitations of psychological testing on the internet. *Journal of Technology in Human Services, 19*, 65–89.
- Bartram, D. (1999). Testing and the Internet: Current realities, issues and future possibilities. *Selection and Development Review, 15*, 3–12.
- 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. (1999). Citizenship and impression management: Good soldiers or good actors? *Academy of Management Review*, *24*, 82–98.
- Bratton, G. R., & Newsted, P. R. (1995). Response effects and computer-administered questionnaires: the role of the entry task and previous computer experience. *Behaviour and Information Technology*, *14*, 300–312.
- Braun, J. R. (1963). Effects of positive and negative faking sets on the survey of interpersonal values. *Psychological Reports*, *13*, 171–173.
- Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research*, *24*, 445–455.
- Cardy, D., & Miller, J. (2003). Technology: Implications for HR. In D. Stone (Ed.), *Advances in human performance and cognitive engineering research* (Vol. 3, pp. 99–118). Amsterdam: JAI.
- Christiansen, N. D., Goffin, R. D., Johnston, N. G., & Rothstein, M. G. (1994). Correcting the 16PF for faking: Effects on criterion-related validity and individual hiring decisions. *Personnel Psychology*, *47*, 847–860.
- Cohen, J. (1988). *Statistical power analysis for the behavioral science* (2nd Ed.). Hillsdale: Erlbaum.
- Dalen, L. H., Stanton, N. A., & Roberts, A. D. (2001). Faking personality questionnaires in personnel selection. *Journal of Management Development*, *20*, 729–741.
- Day, D. V., Unckless, A. L., Schleicher, D. J., & Hiller, N. J. (2002). Self-Monitoring personality at work: A meta-analytic investigation of construct validity. *Journal of Applied Psychology*, *87*, 390–401.
- Dwight, S. A., & Feigelson, M. E. (2000). A quantitative review of the effect of computerized testing on the measurement of social desirability. *Educational and Psychological Measurement*, *60*, 340–360.
- Edens, P. S., & Arthur, W. Jr. (2000). A meta-analysis investigating the susceptibility of self-report inventories to distortion. *Poster presented at the 15th annual meeting of the society for industrial and organizational psychology*. New Orleans, Louisiana.
- 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.
- Evans, D. C., Garcia, D. J., Garcia, D. M., & Baron, R. S. (2003). In the privacy of their own homes: Using the Internet to assess racial bias. *Personality and Social Psychology Bulletin*, *29*, 273–284.
- Furnham, A. (1990). The fakeability of the 16PF, Myers-Briggs and Firo-B personality measures. *Personality and Individual Differences*, *11*, 711–716.
- Gangestad, S. W., & Snyder, M. (2000). Self-Monitoring: Appraisal and Reappraisal. *Psychological Bulletin*, *126*, 530–555.
- Gilmore, D. C., & Ferris, G. R. (1989). The effects of applicant impression management tactics on interviewer judgments. *Journal of Management*, *15*, 557–564.
- Hertel, G., Konradt, U., & Orlikowski, B. (2003). Ziele und Strategien von E-Assessment aus Sicht der psychologischen Personalauswahl. In U. Konradt & W. Sarges (Eds.), *E-Recruitment und E-Assessment* (pp. 37–53). Göttingen: Verlag für Angewandte Psychologie.
- Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology*, *92*, 1270–1285.
- Holden, R. R. (1998). Detecting fakers on a personnel test: Response latencies versus a standard validity scale. *Journal of Social Behavior and Personality*, *13*, 387–398.
- Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection, and amelioration of adverse impact in personnel selection procedures: Issues, evidence, and lessons learned. *International Journal of Selection and Assessment*, *9*, 152–194.
- Hough, L. M., & Schneider, R. J. (1996). Personality traits, taxonomies, and applications in organizations. In K.R. Murphy (Ed.), *Individual differences and behavior in organizations* (pp. 31–88). San Francisco: Jossey-Bass.
- Hoyle, R. H., & Lennox, R. D. (1991). Latent structure of self-monitoring. *Multivariate Behavioral Research*, *26*, 511–540.
- Hsu, L. M., Santelli, J., & Hsu, J. R. (1989). Faking detection validity and incremental validity of the response latencies to MMPI subtle and obvious items. *Journal of Personality Assessment*, *53*, 278–295.
- Hu, L., & Bentler, P. M. (1999). Cutoff criterion for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.
- Joinson, A. N. (2001). Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. *European Journal of Social Psychology*, *31*, 177–192.
- Kelly, D. B., & Greene, R. L. (1989). Detection of faking good on the MMPI in a psychiatric inpatient population. *Psychological Reports*, *65*, 747–750.
- Kleinmann, M., Kuptsch, C., & Köller, O. (1996). Transparency: A necessary requirement for the construct validity of assessment centres. *Applied Psychology: An International Review*, *45*, 67–84.
- Konradt, U., & Hertel, G. (2004). Personalauswahl, Platzierung und Potenzialanalyse mit internetbasierten Verfahren. In G. Hertel & U. Konradt (Hrsg.), *Human Resource Management im Inter- und Intranet* (S. 55–71). Göttingen: Hogrefe.
- Konradt, U., Hertel, G., & Joder, K. (2004). Der Callcenter-Aptitude-Test (CAT). In W. Sarges & H. Wotta-

- wa (Hrsg.). *Handbuch wirtschaftspsychologischer Verfahren* (S. 225–227). Lengerich: Pabst.
- Konradt, U., Hertel, G., & Joder, K. (2003). Web-based assessment of call center agents: Development and validation of a computerized instrument. *International Journal of Assessment and Selection*, *11*, 184–193.
- Konradt, U., Lehmann, K., Böhm-Rupprecht, J., & Hertel, G. (2003). Computer- und internetbasierte Verfahren der Berufseignungsdiagnostik: Ein empirischer Überblick. In U. Konradt & W. Sarges (Eds.). *E-Recruitment und E-Assessment* (pp. 105–124). Göttingen: Verlag für Angewandte Psychologie.
- Lanyon, R. I. (1993). Development of scales to assess specific deception strategies on the Psychological Screening Inventory. *Psychological Assessment*, *5*, 324–329.
- Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, *107*, 34–47.
- Lievens, F., & Harris, M. M. (2003). Research on Internet recruiting and testing: Current status and future directions. In C. L. Cooper & I. T. Robertson (Eds.). *International Review of Industrial and Organizational Psychology* (pp. 131–165). Chichester: Wiley.
- Mahar, D., Henderson, R., & Deane, F. (1997). The effects of computer anxiety, state anxiety, and computer experience on users' performance on computer based tasks. *Personality and Individual Differences*, *22*, 683–692.
- Matheson, K., & Zanna, M. P. (1988). The impact of computer-mediated communication on self-awareness. *Computers in Human Behaviour*, *4*, 221–233.
- Matheson, K., & Zanna, M.P. (1990). Computer-mediated communication: The focus is on me. *Social Science Computer Review*, *8*, 1–12.
- McFarland, L. A., & Ryan, A. M. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology*, *85*, 812–821.
- Meehl, P., & Hathaway, S. (1946). The K Factor as a suppressor in the MMPI. *Journal of Applied Psychology*, *30*, 525–564.
- Mielke, R., & Kilian, R. (1990). Wenn Teilskalen sich nicht zu dem ergänzen, was die Gesamtskala erfassen soll: Untersuchungen zum Self-Monitoring-Konzept. [If subscales do not go together as the concept demands it. Empirical studies of the self-monitoring-concept] *Zeitschrift für Sozialpsychologie*, *21*, 126–135.
- Miller, J.S., & Cardy, R.L. (2000). Self-monitoring and performance appraisal: Rating outcomes in project teams. *Journal of Organizational Behavior*, *21*, 609–626.
- Moon, Y. (1998). Impression Management in computer-based interviews: The effects of input modality, output modality, and distance. *Public Opinion Quarterly*, *62*, 610–622.
- Naglieri, J. A., Drasgow, F., Schmit, M., Handler, L., Prifitera, A., Margolis, A., & Velasquez, R. (2004). Psychological testing on the internet: New problems, old issues. *American Psychologist*, *59*, 150–162.
- Nezlek, J. B., & Leary, M. R. (2002). Individual differences in self-presentational motives in daily social interaction. *Personality and Social Psychology Bulletin*, *28*, 211–223.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw Hill.
- 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.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, *46*, 598–609.
- Paulhus, D. L. (1986). Self-deception and impression management in test responses. In A. Angleitner & J. S. Wiggins (Eds.). *Personality assessment via questionnaires* (pp. 144–165). Berlin: Springer.
- Paulhus, D. L., Bruce, M. N., & Trapnell, P. D. (1995). Effects of self-presentation strategies on personality profiles and their structure. *Personality and Social Psychology Bulletin*, *21*, 100–108.
- Pauls, C. A., & Crost, N. W. (2005). Cognitive ability and self-reported efficacy of self-presentation predict faking on personality measures. *Journal of Individual Differences*, *26*, 194–206.
- Richman, W. L., Kiesler, S., Weisband, S., & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology*, *84*, 754–775.
- Robie, C., Born, M. P., & Schmit, M. J. (2001). Personal and situational determinants of personality responses: A partial reanalysis and reinterpretation of the Schmit et al. (1995) data. *Journal of Business and Psychology*, *16*, 101–117.
- Rosenfeld, P. R., Giacalone, R. A., & Riordan, C. A. (Eds.) (1995). *Impression management in organizations: Theory, measurement, and practice*. New York: Routledge.
- Sackett, P. R., & Lievens, F. (2008). Personnel selection. *Annual Review of Psychology*, *59*, 419–450.
- 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.
- Smith, D. B., & Robie, C. (2004). The implications of impression management for personality research in organizations. In B. Schneider & D. B. Smith (Eds.), *Personality and organizations* (pp. 111–138). Mahwah, NJ: Lawrence Erlbaum Associates.

- Snell, A. F., Sydel, E. J., & Lueke, S. B. (1999). Towards a theory of applicant faking: Integrating studies of deception. *Human Resource Management Review*, *9*, 219–242.
- Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, *30*, 526–537.
- Snyder, M. (1987). *Public appearances-private realities: The psychology of self-monitoring*. New York: W. H. Freeman.
- 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.
- Tabachnick, B. G., & Fidell, L. S. (2006). *Using multivariate statistics* (5 ed.). Boston, MA: Allyn and Bacon.
- Tippins, N., Beaty, J., Drasgow, F., Gibson, W., Pearlman, K., Segall, D., & Shepherd, W. (2006). Unproctored Internet testing in employment settings. *Personnel Psychology*, *59*, 189–225.
- 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.
- Turnley, W. H., & Bolino, M. C. (2001). Achieving desired images while avoiding undesired images: Exploring the role of self-monitoring in impression management. *Journal of Applied Psychology*, *86*, 351–360.
- Van de Vijver, F. J. R., & Harsveld, M. (1994). The incomplete equivalence of the paper-and-pencil and computerized versions of the general Aptitude test Battery. *Journal of Applied Psychology*, *79*, 852–860.
- Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: implications for personality measurement. *Educational and Psychological Measurement*, *59*, 197–210.
- Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, *23*, 3–43.
- Wayne, S. J., & Ferris, G. R. (1990). Influence tactics, affect, and exchange quality in supervisor-subordinate interactions: A laboratory experiment and field study. *Journal of Applied Psychology*, *75*, 487–499.
- Wayne, S. J., & Green, S. A. (1993). The effects of leader member exchange on employee citizenship and impression management behavior. *Human Relations*, *46*, 1431–1440.
- Wayne, S. J., & Liden, R. C. (1995). Effects of impression management on performance ratings: A longitudinal study. *Academy of Management Journal*, *38*, 232–260.
- Zickar, M. J., & Drasgow, F. (1996). Detecting faking on a personality instrument using appropriateness measurement. *Applied Psychological Measurement*, *20*, 71–87.
- Zickar, M. J., & Robie, C. (1999). Modeling faking good on personality items: An item level analysis. *Journal of Applied Psychology*, *84*, 551–563.

Corresponding author:

Prof. Dr. Udo Konradt
 Institute of Psychology
 University of Kiel
 Olshausenstr. 40
 D-24098 Kiel
 GERMANY
 konradt@psychologie.uni-kiel.de

Appendix

Table 1: Means, Standard Deviations of the Personality Scales for Each Condition

Scale	Honest (N = 103)		Fake good (N = 107)		Application (N = 112)		F
	M	SD	M	SD	M	SD	
Willingness to learn	5.28	1.09	5.47	12.5	5.56	1.02	1.70
Social Competence	4.68	1.22	4.81	1.33	4.94	1.09	1.18
Self-motivation	5.15	1.06	5.53	1.10	5.52	0.91	4.74**
Stress resistance	5.29	0.91	5.35	1.07	5.38	0.97	0.26
Conscientiousness	5.70	0.81	5.93	0.83	5.96	0.70	3.61*

Note: * $p < .05$; ** $p < .01$ (two-tailed)

Table 2: Means, Standard Deviations, Reliabilities, and Intercorrelations of the Main Variables

	M	SD	1	2	3	4	5	6	7	8	9	10	11
<i>Personality measures</i>													
1 Willingness to learn	5.43	1.13	(.75)										
2 Social Competence	4.81	1.22	.45**	(.70)									
3 Self-motivation	5.40	1.04	.57**	.42**	(.67)								
4 Stress resistance	5.34	0.98	.62**	.46**	.52**	(.66)							
5 Conscientiousness	5.86	0.79	.46**	.41**	.57*	.49**	(.66)						
<i>Self-Monitoring measures</i>													
6 Social abilities	4.26	1.12	.19**	.27**	.08	.22*	.10	(.68)					
7 Inconsistency	2.88	1.09	-.07	-.08	-.13*	-.16**	-.24**	.12*	(.72)				
8 Social comparing	4.48	1.16	-.08	-.03	-.13*	-.08	-.03	-.05	.18**	(.70)			
<i>Socio-demographic variables</i>													
9 Age	29.61	5.89	.15**	.17**	.25**	.21**	.20**	-.17*	-.06	-.01	n. a.		
10 Gender ^a	--	--	.19**	.07	.06	.07	-.00	-.04	.01	-.00	.00	n. a.	
11 Education ^b	--	--	.13*	.09	-.01	.11	-.08	.12*	.05	.07	-.05	-.06	n. a.

Note: N = 312. a: Dummy coded variable (1 = male, 2 = female). b: 1 = low level of education, 4 = high level of education. Numbers on the diagonal are actual alpha coefficients. * $p < .01$ (two-tailed). The above values were computed across conditions.

Table 3: Hierarchical Regression Analyses Predicting the Overall Personality Measure

Step	Variables	<i>B</i>	<i>Beta</i>	<i>Multiple R</i>	<i>Adj R²</i>	<i>Delta R²</i>	<i>F</i>
Honest condition							
1	Age	.01	.06	.02	-.01	.02	.076
	Education	.08	.11				
2	Social abilities	.34***	.48	.30	.28	.28	13.00***
	Inconsistency	-.16	-.22				
	Social comparing	-.09	-.13				
Fake good condition							
1	Age	.05***	.33	.12	.11	.12	7.22***
	Education	.01	.12				
2	Social abilities	.15*	.20	.17	.13	.05	2.12
	Inconsistency	-.16*	-.18				
	Social comparing	.01	.02				
Application condition							
1	Age	.05***	.39	.11	.10	.11	6.41*
	Education	-.03	-.05				
2	Social abilities	.21***	.32	.25	.21	.14	5.94***
	Inconsistency	-.12*	-.21				
	Social comparing	.01	.01				

Note: * $p < .05$, ** $p < .01$, *** $p < .001$. Hypthesized effects are one-tailed tests and all other results represent two-tailed tests.