HEXACO Personality Predicts Counterproductive Work Behavior and Organizational

Citizenship Behavior in Low-Stakes and Job Applicant Contexts

Jeromy Anglim
Deakin University (1)

Filip Lievens
Singapore Management University (2)

Lisa Everton
Deakin University (1)

Sharon L. Grant
Swinburne University (3)

Andrew Marty
SACS Consulting (4)

Author Note

1 School of Psychology, Deakin University, Geelong, Australia; 2. Singapore Management University, Lee Kong Chian School of Business, Singapore; 3. Department of Psychological Sciences, Swinburne University of Technology 4. Andrew Marty, SACS Consulting, Melbourne, Australia. Declarations of interest: The fifth author (AM) uses the HEXACO model of personality in a consulting context. The other authors have no interests to declare. Contribution note: JA, LE, and AM were involved in study design and data collection. JA performed the data analysis. All authors contributed to writing of the manuscript. Correspondence concerning this article should be addressed to Jeromy Anglim, School of Psychology, Deakin University, Locked Bag 20000, Geelong, 3220 Australia. Email: jeromy.anglim@deakin.edu.au

Abstract

This study examined the degree to which the predictive validity of personality declines in job applicant settings. Participants completed the 200-item HEXACO Personality Inventory-Revised, either as part of confidential research (347 non-applicants) or an actual job application (260 job applicants). Approximately 18-months later, participants completed a confidential survey measuring organizational citizenship behavior (OCB) and counterproductive work behavior (CWB). There was evidence for a small drop in predictive validity among job applicants, however honesty-humility, extraversion, agreeableness, and conscientiousness predicted lower levels of CWB and higher levels of OCB in both job applicants and non-applicants. The study also informs the use of the HEXACO model of personality in selection settings, reporting typical levels of applicant faking and facet-level predictive validity.

Keywords: faking, employee selection, HEXACO, organizational citizenship behavior, counterproductive work behavior

1. Introduction

Personality testing for employee selection offers the potential to identify job applicants who will contribute positively to workplace culture and abstain from damaging behaviors such as fraud, theft, and harassment. However, job applicants are understandably motivated to make a positive impression on employers and tend to engage in response distortion when completing personality tests (Birkeland, Manson, Kisamore, Brannick, & Smith, 2006). This raises the question of whether job applicant response distortion reduces the predictive validity of personality measures in high stakes settings (Komar, Brown, Komar, & Robie, 2008; Morgeson et al., 2007a, 2007b; Rothstein & Goffin, 2006).

Thus, the primary aim of the present study was to examine the effect of the job applicant context on the predictive validity of personality tests. A secondary aim was to examine the predictive validity of the broad and narrow traits of the HEXACO model of personality. To achieve these aims, we utilized two large existing databases of applicants and non-applicants that included responses for the HEXACO Personality Inventory - Revised. We engaged in follow-up measurement of these samples to measure their organizational citizenship behavior and counterproductive work behavior in a low-stakes confidential research setting. As will be discussed, this design overcomes some of the limitations of past literature on comparative predictive validity. In addition to the primary aim, the data also allowed for an examination of the predictive validity of the HEXACO model of personality, which is an increasingly popular alternative to the Big Five (Lee, Ashton, Morrison, Cordery, & Dunlop, 2008), particularly as it relates to unethical and deviant behavior (for a review see, Ashton, Lee, & de Vries, 2014). As we had data on the full 200-item measure of the HEXACO PI-R that provides reliable measurement of personality facets, the study also contributes to discussion about the relative merits of narrow traits in employee selection (Anglim, Bozic, Little, &

Lievens, 2018; Ashton, 1998; Christiansen & Robie, 2011; Ones & Viswesvaran, 1996; Salgado, Moscoso, & Berges, 2013; Tett, Steele, & Beauregard, 2003).

1.1.   The Job Applicant Context and Predictive Validity

The relationship between personality and employee behavior has generated considerable research in I/O psychology (for a review, see Barrick & Mount, 2012) with numerous meta-analyses that synthesize correlations between the Big Five and work performance (Barrick & Mount, 1991; Barrick, Mount, & Judge, 2001; Hurtz & Donovan, 2000; Salgado, 1997; Tett, Jackson, & Rothstein, 1991). More recently, meta-analyses have focused on contextual performance outcomes such as organizational citizenship behavior (OCB), prosocial behavior which can benefit the organization or people within it, and counterproductive work behavior (CWB), behavior which can harm the organization or people within it. A meta-analysis of the correlations between self-rated Big Five personality and OCB obtained small but significant mean raw correlations for conscientiousness, agreeableness, emotional stability, extraversion and openness (Chiaburu, Oh, Berry, Li, & Gardner, 2011). With regard to CWB, a meta-analysis by Berry, Ones, and Sackett (2007) showed that low agreeableness and low conscientiousness were the traits most strongly related to overall CWB.

Researchers have also questioned whether the predictive validity of personality is retained when respondents are motivated and able to fake (Rothstein & Goffin, 2006). Respondents generally recognize which responses to personality test items are more socially desirable and do in fact respond in more socially desirable ways in job applicant settings (for a review, see Rothstein & Goffin, 2006). A meta-analysis by Birkeland et al. (2006) found that personality test scores were approximately half a standard deviation higher on conscientiousness and emotional stability in job applicant samples. A meta-analysis of instructed faking studies found average changes of around three-quarters of a standard deviation (Ones, Viswesvaran, & Schmidt, 1993). In addition to studies of instructed faking on HEXACO personality in the laboratory environment (Grieve & De Groot, 2011; MacCann, 2013), a recent study by Anglim, Morse, De Vries, MacCann, and Marty (2017) compared a large samples of job applicants and non-applicants on the HEXACO-PI-R and found that job applicants scored higher on honesty-humility, extraversion, agreeableness, and conscientiousness.

Although there is relative consensus that response distortion occurs in applicant settings, there is less agreement whether such settings also reduce the predictive validity of personality. Rothstein and Goffin (2006) concluded that while faking reduces the validity of personality testing, the predictive validity that remains is still sufficient to justify the use of personality testing. In contrast, Morgeson et al. (2007a) suggested that the evidence that the applicant context lowers predictive validity is inconclusive. They noted that relatively large samples are required to determine whether differences in predictive validity are significant, given the small observed correlations between personality and performance. Given that several studies have found that applicant personality predicts job performance (Barrick & Mount, 1996), a key issue is quantifying how much, if at all, predictive validity is reduced by applicant faking.

A range of empirical approaches have been taken to examine changes in predictive validity, and each has its limitations. First, early research showed that applying adjustments for faking based on impression management scales did not increase predictive validity (Barrick & Mount, 1996; Christiansen, Goffin, Johnston, & Rothstein, 1994; Ones, Viswesvaran, & Reiss, 1996; Schmitt & Oswald, 2006). However, faking is better understood as the difference between a latent true score and the latent faked score (Peterson, Griffith, Isaacson, O'Connell, & Mangos, 2011). Impression management scales are contaminated with substantive variance and are not direct measures of the amount of faking, and therefore cannot be used to examine how predictive validity changes with faking (Anglim et al., 2017; De Vries, Zettler, & Hilbig, 2014; Mueller-Hanson, Heggestad, & Thornton III, 2003; Uziel, 2010).

Second, several meta-analyses comparing the predictive validity of job applicant with non-applicant samples have shown that personality is predictive of performance outcomes in job applicant contexts. For example, Schmitt, Gooding, Noe, and Kirsch (1984) found average correlations between personality and supervisor performance ratings of .34 for concurrent designs using existing employees, and .30 for the predictive designs using applicants. Hough's (1997) meta-analysis indicated that predictive validities of personality were on average .07 higher in incumbent samples compared to applicant samples. In contrast, Tett et al. (1991) indicated that recruitment samples had a larger sample weighted validity correlation than incumbents (mean r of .20 versus .14), although 80% of the recruitment sample came from only one sample, the military Project A dataset (Campbell, 1990).

Finally, studies that have examined predictive validity by comparing applicants and non-applicants in a single study have had methodological limitations. Some research used laboratory designs (e.g., Douglas, McDaniel, & Snell, 1996; Mueller-Hanson et al., 2003) where participants either answer a personality test honestly or under some form of instructed faking conditions. These studies often used university samples and used grade point average as an outcome for evaluating predictive validity. While job applicant role plays are a useful tool, researchers are generally most interested in how these results generalize to high-stakes settings, and for that, getting the responses of actual job applicants is preferred. Other studies compared predictive validity in actual job applicants with non-applicants such as incumbents (Anglim et al., 2018; Ellingson, Sackett, & Connelly, 2007; Hough, Eaton, Dunnette, Kamp, & McCloy, 1990; Jeong, Christiansen, Robie, Kung, & Kinney, 2017; Mueller-Hanson et al., 2003). For example, Jeong et al. (2017) found that personality was less predictive of supervisor ratings in hired applicants than incumbents. Some studies comparing applicants and non-applicants have design confounds that compromise inferences about effects of the applicant context. For example, there has often been a time delay between applicant personality measurement and outcome measurement (i.e., predictive design) but no time delay for the non-applicants (i.e., incumbents, concurrent design). Differential range restriction between groups and other issues of group comparability represent additional challenges. Furthermore, many previous studies have had modest sample sizes and research on differential

predictive validity across contexts benefits from large sample sizes. Thus, further research with larger samples that controls for these confounds is still needed.

1.2. The HEXACO Model of Personality

Historically, personality research began with a proliferation of traits followed by the widespread acceptance that there were five broad personality traits (Costa & MacCrae, 1992; Goldberg, 1993). In addition to higher-order models (Digman, 1997; Musek, 2007; Van der Linden, te Nijenhuis, & Bakker, 2010) and competing models of narrow facets (DeYoung, Quilty, & Peterson, 2007; Paunonen & Ashton, 2001; Paunonen & Jackson, 2000; Roberts, Chernyshenko, Stark, & Goldberg, 2005; Salgado et al., 2015), one of the biggest challenges to the Big Five comes from models that propose alternative broad traits (for a review, see Anglim & O'Connor, 2018). Prominently, Lee, Ashton, and colleagues developed the six factor HEXACO model based on extensive lexical studies in a range of languages (Ashton et al., 2014; Lee & Ashton, 2008). The HEXACO model largely retains extraversion, openness, and conscientiousness from the Big Five, but Big Five agreeableness and neuroticism are reconfigured and expanded into HEXACO emotionality, agreeableness, and honesty-humility. Initial studies using HEXACO personality to predict CWB have tended to identify Honesty–humility as a comparatively strong predictor of lower levels of CWB and other deviant behaviors (Ceschi, Sartori, Dickert, & Costantini, 2016; Chirumbolo, 2015; de Vries & van Gelder, 2015; Lee, Ashton, & de Vries, 2005; Marcus, Lee, & Ashton, 2007; Oh, Lee, Ashton, & De Vries, 2011).

Other research has examined whether narrow facets provide additional benefits in predicting workplace behavior (Anglim et al., 2018; Anglim & Grant, 2014; Cronbach & Gleser, 1965; Ones & Viswesvaran, 1996). A review of 11 empirical studies examining prediction of facets (e.g, Jenkins & Griffith, 2004; Tett et al., 2003; Vinchur, Schippmann, Switzer III, & Roth, 1998) and traits outside the Big Five (e.g., Ashton, 1998; Conte & Gintoft, 2005; Crant, 1995; Lounsbury, Gibson, & Hamrick, 2004) for job performance indicated that narrow facets either correlated more with performance than factors or provided significant incremental prediction over factors (Rothstein & Goffin, 2006). In addition, several more recent studies examined incremental prediction of facets (de Vries, de Vries, & Born, 2011; Jenkins & Griffith, 2004; Judge, Rodell, Klinger, Simon, & Crawford, 2013; Salgado et al., 2013; Salgado et al., 2015; Tett et al., 2003; Vinchur et al., 1998; Ziegler et al., 2014) and novel traits outside the five-factor model (Ashton, 1998; Conte & Gintoft, 2005; Crant, 1995; Lounsbury et al., 2004). Many of these studies suggest that inclusion of facets can increase prediction by about a third to a half. Nonetheless, existing research examining correlations between personality facets and OCB/CWB has rarely used the HEXACO model, and has instead used other personality frameworks such as the NEO-PI-R. Given the increased practitioner interest in the HEXACO framework, research is needed to understand how the narrow facets of the HEXACO model predict OCB and CWB. In particular, such research can guide decisions about how employers should weight the broad and narrow traits of the HEXACO model in selection systems. In may also enable the development of shorter and more efficient measures. From a theoretical

perspective, understanding the relative importance of facets can provide a richer understanding of how predictive validity operates at the level of broad traits by highlighting particular narrow facets that are more or less important.

1.3. The Current Study

On the backdrop of the above, the main aim of the current study was to examine how the predictive validity of personality tests varies between job applicant and non-applicant settings. A secondary aim was to examine the predictive validity of HEXACO personality for OCB and CWB. Respondents were either job applicants or research participants who had completed the 200-item HEXACO-PI-R with six factors and 25 facets (for information on this large baseline sample, see Anglim et al., 2017). Respondents also completed a follow-up survey, which included measures of OCB and CWB, in a low-stakes testing environment (confidential research), approximately 18 months after initial testing. By having an extended period between initial personality assessment and the follow-up survey, we were able to ensure that applicants and non-applicants answered follow-up questions about OCB and CWB equally honestly. The design of the current study also overcomes some limitations with previous research. In particular, much previous research confounds the setting (applicant versus non-applicant) with the design (predictive versus concurrent). For example, non-applicants are commonly incumbents where outcomes are measured concurrently. In our design, all outcomes are measured in a time-lagged fashion.

2. Method

2.1. Participants and Procedure

The sample consisted of 607 participants (347 non-applicants; 260 job applicants). Note that we use the terms "non-applicants" and "job applicants" to distinguish how baseline personality was measured, even though both groups completed follow-up measurement in a non-applicant context. The time between original and follow-up measurement was generally over one year and similar for applicants ($M = 1.6$ years, $SD = 1.1$, range: 0.04–4.06) and non-applicants ($M = 1.9$ years, $SD = 1.0$, range: 0.1–3.81). Applicants were 31% male and non-applicants were 60% male. At the time of follow-up, applicants were aged between 19 and 65 years ($M = 41.2$, $SD = 12.1$) and non-applicants were aged between 22 and 72 years ($M = 50.7$, $SD = 10.4$). Both samples were well educated, with most having a bachelor's degree or higher (applicants 68%; non-applicants 78%). Almost all participants lived in Australia (applicants 98%; non-applicants 96%), and most were born in Australia (applicants 79%; non-applicants 72%). Most were working full-time (applicants 71%; non-applicants 69%). Participants worked in a broad cross-section of industries with similar characteristics across applicants and non-applicants with the most common industries being health and community services (24%), government administration and defense (17%), education (9%), property and business services (5%), and finance and insurance (5%). Overall, the non-applicants had a slightly shorter gap between baseline and follow-up measurement than the applicants and were significantly older and more likely to be male. We

present several analyses in the results section showing that controlling for age, gender, and time between sessions had minimal effect on results. The study received ethics approval from the first author's University, Faculty of Health Human Research Ethics committee.

Participants were obtained from the database of an Australian human resources consulting organization. All respondents had completed the 200-item version of the HEXACO-PI-R either as part of confidential, low-stakes internal research ('non-applicants') or as part of a job application ('applicants'). Non-applicants had completed the personality questionnaire in one of several discrete data collection periods from mid-2010 to early-2015. Applicants had completed the personality questionnaire, over the same period, while applying for jobs at organizations that were using the consulting organization's psychometric testing services. Analyses of differences on this larger sample of HEXACO-PI-R are reported in Anglim et al. (2017).

In August of 2015, both job applicants and non-applicants were sent an email inviting them to complete a questionnaire on employee behavior, attitudes, and personality. Prospective participants were informed that the survey was for research purposes only and all responses were confidential. They were also informed that by completing the survey and providing their email address that they would enter a draw to win a travel voucher of AUD$3,000 or one of three shopping gift cards valued at AUD$500, AUD$200 or AUD$100. Participants consented to allow their previous personality data to be matched with their responses for the additional measures via their email address. Participants completed measures of OCB, CWB, demographics, an additional measure of HEXACO personality (the SACS-6), and other measures not used in the present study. Of the 776 participants who provided consent for data matching, 611 participants matched an existing record in the HEXACO-PI-R database, and four were excluded because they were missing more than 15% of the outcome data. There was no missing data on the HEXACO-PI-R.

The sample size in the present study was determined by the response rate of the underlying applicant and non-applicant baseline personality databases. The unique design limited the potential to simply recruit more participants. The current study is concerned with quantifying a range of different parameters, each with different expected effect sizes. For example, statistical power was approximately 68%, 96%, for group differences of .2 and .3 standard deviations respectively, and 100% for "medium" effects and above (e.g., d > 0.5), commonly seen in the literature (Birkeland et al., 2006). In terms of quantifying group differences in correlations, assuming a $r = .20$ in the applicant sample and $r = .40$ in the non-applicants, the study has 76% power. Thus, the study is able to identify if the applicant context leads to substantial reductions in predictive validity. However, for more subtle reductions in predictive validity (e.g., correlations changing from .20 to .15, etc.) larger sample sizes are needed, and this study contributes data that can inform meta-analytic estimates.

2.2.   Materials

All scales were scored as the mean after any necessary item reversal. Scale reliabilities are presented in the Results section.

2.2.1.   Baseline Personality

The 200-item HEXACO Personality Inventory Revised (Lee & Ashton, 2004) was used to measure personality. The HEXACO-PI-R measures six domains of personality (honesty-humility, emotionality, extraversion, agreeableness, conscientiousness, and openness). Each domain is composed of four underlying facets (8 items per facet), and there is one additional interstitial facet called altruism. Participants indicated their agreement with each statement on a 5-point scale from 1 = *strongly agree* to 5 = *strongly agree*.

2.2.2.   Follow-up Personality

The six HEXACO domains were measured at follow-up using the 253 item version of the SACS-6, developed by SACS Consulting to measure the HEXACO factors and facets. Although items are generally worded to be more work relevant, the correlation between corresponding domain scales for the HEXACO-PI-R are high (correlations in the non-applicant sample were .72, .75, .73, .69, .67 and .50 for honesty, emotionality, extraversion, agreeableness, conscientiousness, and openness respectively). These correlations are similar to those observed in test-retest studies examining intervals over one year using NEO-PI domains that typically obtain correlations in the .70 to .85 range (Costa & MacCrae, 1992). We used the scores on the SACS-6 to see whether applicants and non-applicants answered similarly on personality at follow-up, where the absence of group differences in personality were used as additional evidence that baseline personality differences were due to the job applicant setting as opposed to underlying differences between the groups.

2.2.3.   Organizational Citizenship Behavior (OCB)

OCB was measured using Lee and Allen's (2002) Organizational Citizenship Behavior Scale. This 16-item scale includes two subscales measuring interpersonal and organizational directed organizational citizenship behaviors. Participants rated how often they engage in certain behaviors on a 7-point scale from 1 = *never* to 7 = *always*.

2.2.4.   Counterproductive Work Behavior (CWB)

CWB was assessed using the Workplace Deviance Scale (Bennett & Robinson, 2000). The scale includes 19 items and consists of two subscales (interpersonal and organizational). Items were answered on a 7-point scale where 1 = *never*, 4 = *several times a year*, and 7 = *daily*.

3.   Results

3.1.   Preliminary Validity Checks

Before examining differences in predictive validity between job applicants and non-applicants and to attribute differences in predictive validity to the effect of the job applicant context, we verified whether responses by job applicants at follow-up were of approximately equivalent honesty as those of non-applicants. Applicants and non-applicants showed minimal differences on personality at follow-up, where mean absolute Cohen's d was .14 at follow-up compared to .50 at baseline. There were also no significant differences between applicants and non-applicants in CWB and OCB. Thus, it seems reasonable to assume that applicants and non-applicants answered the follow-up survey with roughly equivalent levels of honesty.

A second validity issue is whether differences in applicants and non-applicants at baseline can be attributed to the job applicant context as opposed to other differences between groups. In general, the absence of differences in follow-up personality can also be used as evidence that the differences observed at baseline are attributable to the applicant context. Furthermore, we performed a range of sensitivity analyses that showed that group differences in baseline personality exhibited minimal changes when controlling for gender or age (see online supplement for details), although when controlling for gender, applicants reported being slightly lower on emotionality (see also, Anglim et al., 2017).

3.2.   Job Applicant Response Distortion

To assess the magnitude and nature of job applicant response distortion, we examined group differences in means, standard deviations, factor structure, and stability (see online supplement for further details). First, in terms of scale means (see Table 1 for domain-level differences; and supplemental Table S1 for facet-level differences), applicants showed moderate to high levels of response distortion with applicants scoring between two-thirds and one standard deviation higher on honesty-humility [$d = 0.68$, 95% $t(605) = 9.19$, $p <.001$], extraversion [$d = 0.64$, $t(605) = 8.78$, $p <.001$], agreeableness [$d =0.94$, $t(605) = 12.39$, $p <.001$] and conscientiousness [$d = 0.65$, $t(605) = 8.61$, $p <.001$] compared to non-applicants; 95% confidence intervals for d values are reported in table and calculated using the psych package in R (Revelle, 2018). No significant differences were found for emotionality and openness.

Second, standard deviations for personality scales in the applicant sample were on average only 80% of those for non-applicants. Only about a quarter of this reduction could be attributed to scale means in applicants being closer to scale end points.

Third, in terms of factor structure, loading of the 25 facets when extracting six factors in an exploratory factor analysis showed good correspondence to theorized structure in both applicant and non-applicants. Importantly, the size of the first unrotated principal component for facets was larger for applicants (27%) than non-applicants (21%). For a detailed item-level confirmatory factor analytic

examination of the HEXACO-PI-R of applicants and non-applicants, drawn from the larger sample from which the current sample was drawn, see Anglim et al. (2017).

Finally, there was some evidence that baseline personality predicted follow-up personality better in the non-applicant sample. Correlations between baseline and follow-up for corresponding domains were larger on average for non-applicants (mean $r = .68$) than applicants (mean $r = .54$). In addition, six regression models were estimated predicting each of the follow-up personality domains from applicant context, the six baseline personality measures and the interaction between context and the baseline measures. The standard deviation of the residuals were on average 8% larger for applicants than non-applicants.

Overall, these results are consistent with a moderate amount of response distortion. The reduction in standard deviations suggests that participants converge in their responses. In particular, it is consistent with a process whereby those who are truly lower on a socially desirable trait will change their responses more than those whose true responses are already relatively socially desirable. The larger first factor suggests that applicants focus somewhat more on the social desirability of items than on the substantive content when responding. The reduced prediction of "honest" follow-up data reinforces the point that an additional source of variance is influencing responses. Nonetheless, there is still good retention of factor structure and good prediction of follow-up responses in applicants.

3.3. Predictive Validity in Applicants and Non-Applicants

Table 1 presents correlations of personality with OCB and CWB in applicants and non-applicants, and Table 2 shows the results of regression models predicting OCB and CWB from HEXACO domain scores for applicants and non-applicants (see online supplement for complete correlation matrices for applicants, non-applicants, and combined data). Before discussing the effect of the applicant context, we first focus on the general predictive validity of personality in non-applicants. Honesty-humility, extraversion, agreeableness, and conscientiousness showed significant negative correlations with CWB and positive correlations with OCB (absolute correlations ranging from .17 to .39). Correlations with OCB and CWB for emotionality and openness were closer to zero, with the only significant correlation being for openness and OCB ($r = .13$). In general, honesty-humility was relatively more related to CWB than OCB, and extraversion was relatively more related to OCB than CWB. Adjusted r-squared values for regression models in non-applicants with the six personality factors as predictors explained 18% of variance for CWB and 21% of variance in OCB.

We then performed various analyses to assess the effect of applicant context on predictive validity. First, correlations of HEXACO personality with outcomes were fairly similar for applicants and non-applicants. Taking correlations that are above .20 in either context, the mean absolute correlation was .052 larger in non-applicants compared to applicants. Only one correlation was significantly different at the .01 level (see supplement for details): extraversion and CWB was more strongly related in non-applicants ($r = -.09$ for non-applicants versus; $r = -.30$ for non-applicants). Second, personality predicted approximately 50% more variance

in outcomes for non-applicants than for applicants, although this difference was not statistically significant at the .05 level. Thus, while the trend in the data supports greater incremental prediction in non-applicants, the null hypothesis of equal variance explained in applicants and non-applicants was not rejected. Third, we examined the standard error of the regression models. Standard errors provide an unstandardized test of prediction and larger standard errors indicate poorer prediction. Regression models indicated that applicants slightly larger standard errors for CWB (0.57 for applicants versus 0.53 for non-applicants), and what smaller standard errors for OCB (0.62 for applicants versus 0.78 for non-applicants). This difference for OCB seemed to largely reflect the slightly larger standard deviations in OCB for non-applicants (SD = 0.87 versus 0.67).

To further examine the effect of context, we fit regression models predicting each outcome from personality and context, and then a second model including the context by personality interaction for each personality variable. There were no significant interactions between personality and context for OCB. However, there were significant interactions for CWB. The interaction model had adjusted R-squared of .150 compared to .135 for the model without interactions, and this increase was statistically significant, $F(6, 593) = 2.71$, $p = .01$. There were significant interactions of context with honesty-humility ($p = .01$) and extraversion ($p = .03$). Specifically, honesty-humility had a stronger coefficient in applicants (-0.45) than non-applicants (-.14) and the coefficient for extraversion was 0.10 in applicants and -0.19 in non-applicants.

Further analyses examined the robustness of these regression models when age and gender were included (see online supplement for regression coefficients). Results were largely unaltered for OCB, but were somewhat different for CWB. Specifically, regression coefficients indicated that age predicted lower levels of CWB, and in applicants being male predicted higher levels of CWB. When age and gender were included adjusted r-squared (non-applicants .19; applicants .18) and standard errors (non-applicants .53, and applicants .54) were quite similar. When included in a combined regression model, the context interactions for age ($p = .27$) and gender ($p = .07$) were not statistically significant. Similarly, the context interactions for honesty-humility ($p = .13$) and extraversion ($p = .14$) were no longer statistically significant.

An additional analysis examined the predictive validity of baseline personality after controlling for follow-up personality scores. Specifically, residualized baseline personality scores were generated by predicting each of the six baseline personality scores from context and the six follow-up personality scores. A regression was then run predicting OCB and CWB from residualized baseline personality, context, and the interaction between context and residualized baseline personality (see Table 3). If baseline personality was merely another measure of personality, then the coefficients should be close to zero or broadly in the same direction as the unresidualized regression but smaller on the basis that repeated measurement should increase reliability of measurement slightly. Alternatively, for applicants, in addition to measurement error, it should also index the amount of response distortion. If response distortion improves predictive validity, then coefficients should be larger in applicants, particularly on traits that

are normally predictive (e.g., conscientiousness). If response distortion merely adds noise or does not damage predictive validity, then the personality coefficients for applicants should be close to zero. Finally, if response distortion in and of itself predicts more CWBs or OCBs then we would expect to see coefficients in the applicants that are the opposite sign of the non-residualized coefficients in non-applicants. In general, Table 3 suggests that residualized scores provide limited prediction of CWB and OCB. A couple of coefficients are statistically significant (e.g., standardized beta of -.16 for residualized emotionality in non-applicants, and .12 for residualized conscientiousness in non-applicants), but these are fairly small and might be type 1 errors. None of the context by personality interaction effects were statistically significant, although it should be noted that the coefficient for conscientiousness in non-applicants is in the same direction for non-applicants but is basically zero for applicants.

3.4. Facet-Level Analysis

Finally, to provide a more complete understanding of how HEXACO personality predicted OCB and CWB, we report fact-level correlations and regression models. On average, facets explained approximately a third more variance than domains for both applicants and non-applicants. To assess the precision of these estimates in an unbiased way, double-adjusted-bootstrap standard errors for change in estimated population variance explained by facet and domain regression models were computed (Anglim & Grant, 2014, 2016). Specifically, adjusted r-squared for domains and facets respectively was as follows: non-applicant CWB (domains .18, facets .21, delta = .04, 95% CI [0, .09]), non-applicant OCB (domain .21, facets .32, delta = .11, 95% CI [.05, .18]), applicant CWB (domains .11, facets .19, delta = .08, 95% CI [0, .18]), applicant OCB (domains .13, facets .16, delta = .03, 95% CI [0, .12]).

To examine which facets were contributing to incremental prediction, we calculated zero-order correlations between HEXACO facets and outcomes and semi-partial correlations between HEXACO facets and outcomes where HEXACO facets were adjusted for overlap with the six HEXACO domains (see Table 4) (for further details on facet residualization, see Anglim & Grant, 2014; Anglim & O'Connor, 2018). While there are many significant zero-order correlations, the semi-partial correlation matrix presents a more parsimonious representation of what is unique about the facets over and above HEXACO domains. In particular, in the non-applicant context, lower fearfulness and higher diligence and altruism correlated over and above domain scores with OCB. The applicant sample had only one significant semi-partial correlation and this differed from the non-applicant sample (unconventionality with CWB).

4. Discussion

The current study assessed the degree to which the predictive validity of personality is attenuated when measured in a job applicant context, and provided a general assessment of the ability of the HEXACO model of personality to predict OCB and CWB. Results showed that when personality measures were completed for employee selection purposes, applicants responded in a more socially desirable

manner. This was reflected in elevated means, smaller standard deviations, smaller correlations with follow-up personality measurement, and larger traits intercorrelations. Despite this response distortion, only a small drop in the predictive validity of personality was observed in applicants compared with non-applicants, and the size of the reduction varied based on how predictive validity was operationalized. Sample estimates broadly suggested a reduction in predictive validity of around a third for the job applicant group. More generally, the four traits of honesty-humility, extraversion, conscientiousness and agreeableness correlated positively with OCB and negatively with CWB. Honesty-humility was particularly important for CWB.

4.1. Predictive Validity in the Applicant Context

Reflecting the results in Anglim et al. (2017), the applicant context had a number of effects on responses to the HEXACO measure of personality. First, applicant scale means were notably higher for honesty-humility, extraversion, agreeableness, and conscientiousness and broadly consistent with other meta-analytic estimates of real-world applicant faking (Birkeland et al., 2006). Second, the standard deviations for scales were substantially reduced. This is consistent with a process whereby honest responses that are further away from a socially desirable ideal are more likely to move towards that ideal. Third, while the applicant context had minimal effect on the pattern of factor loadings, it did increase the size of correlations between items, facets, and factors and the proportion of variance explained by the first component. This increase in variance in the first factor may be evidence that there are individual differences in the amount of faking which in turn increases the likelihood that predictive validity of personality tests will be reduced in the non-applicant setting. Fourth, correlations with follow-up personality were slightly lower in applicants compared to non-applicants.

Nonetheless, there is some evidence that the change in rank ordering of applicants as a result of response distortion may be meaningful but modest. There were still strong correlations between personality scores in applicants and follow-up low-stakes scores. Ultimately, it is individual differences in the amount of response distortion that changes the rank ordering of applicants, and the reduction in standard deviation suggests that greater response distortion occurs for those who are honestly far from the social ideal (e.g., truly low on conscientiousness). This then leads to a compression of scores, but doesn't change the rank ordering as much as would occur if the variation in faking was unrelated to true scores. Assuming that this modest change in rank ordering was mostly noise, then this would suggest that only a small reduction in predictive validity would result.

Consistent with response distortion introducing some noise in measurement, results generally suggested a small drop in predictive validity when personality tests are administered in an employee selection context. Correlations were somewhat lower in non-applicants, but only the extraversion–CWB correlation was significantly reduced. Variance explained by regression models also appeared lower in applicants. That said, even with relatively large sample sizes, the 95% confidence interval on these differences remain relatively large. The

results also highlight the importance of considering which index of prediction should be used to compare predictive validity. In particular, standardized measures of effect (e.g., correlations, standardized coefficients) are lower when standard deviations of either predictors or outcomes are reduced even when the regression equation is unchanged; and the applicant context leads to lower standard deviations in personality scores. Despite the inherent challenge posed in estimating differential prediction, the study contributes an important data point for this estimate. Results are consistent with some meta-analytic research (Hough, 1997; Tett et al., 1991) that suggests that applicant samples have slightly lower predictive validity. It is also broadly consistent with the simulation study by Komar et al. (2008) if we assume that faking mostly contributes noise, faking is moderate in magnitude, and there is a modest amount of individual differences in the magnitude of faking.

There was also some evidence that the regression equation for predicting OCB and CWB is altered in selection settings. Honesty-humility appeared more important and extraversion less important when predicting CWB in applicants. Given the number of possible context by trait interactions and that the interactions effects were no longer statistically significant after controlling for demographics, these findings should be treated as tentative. If the findings do prove to replicate, one interpretation is that extraversion in the HEXACO model loads very highly on a general evaluative factor. Thus, it may be that this factor is less robust in the context of participant response distortion, whereas more descriptive and directly relevant factors to CWB like honesty-humility retain their predictive validity. This also reinforces the importance of using applicant personality norms when reporting standardized scores on applicants, and all else being equal, using applicant validation studies to derive weightings for employee selection algorithms.

4.2. Predictive Validity of HEXACO Personality

Overall, HEXACO personality was a good predictor of OCB and CWB. Conscientiousness, extraversion, and agreeableness all emerged as moderate to strong predictors of OCB and CWB. A particular benefit of the HEXACO honesty-humility factor was in predicting counterproductive work behavior. The interstitial facet of altruism also emerged as a reasonably good predictor. It should be noted that because outcome measures were typically obtained a year or more after personality assessment, correlations may be slightly lower than expected in cross-sectional studies. However, given the stability of personality over time and the fact that time between testing did not moderate the association between personality and outcomes, such attenuation is likely to be mild. In addition, organizations are specifically interested in how well personality tests administered during the selection process predict job behavior over the life of the employee.

The study also examined incremental prediction by narrow traits over broad factors using the six domains and 25 facets of the HEXACO model. While including facets led to a meaningful increase in population variance explained of about a third, questions remain about the robustness of facet-level equations in the current sample. Only a few facets provided incremental prediction over the six HEXACO domains at the .001 level and to the extent that the applicant sample can

be seen as a replication, the samples did not identify the same facets. Given the size of the incremental variance explained by facets in the regression model but the paucity of significant semi-partial correlations, it may be that the contribution of facets arises from many small incremental predictions that individually are not large enough to be deemed statistically significant. The smaller ratio of facets to factors (25 / 6) in the HEXACO model compared to some other models (e.g., 30 / 5 in the NEO) may also lead to fewer facets being identified as incremental predictors. Even larger validation samples are likely to be required in order to obtain a regression equation with facets that yield superior cross-validated prediction compared to a model with just broad traits.

4.3. Limitations

Several limitations should be noted. First, the study used self-report measures of OCB and CWB. Nonetheless, some research suggests that self-reports particularly in relation to CWB provide unique benefits and may be as valid as other-ratings (Berry, Carpenter, & Barratt, 2012). Also, given the focus of this study was on the relative prediction of personality in applicant and non-applicant contexts, the results presented still speak to the size of comparative validity. Second, beyond the usual concerns about common method variance, it is worth considering whether there is anything particular about an "honest-mindset" that may bias validities upwards in the non-applicant context relative to the applicant context. Specifically, an anonymous reviewer suggested that common-method variance may be amplified to a greater extent in the nonapplicant-nonapplicant participants because both the situation and the source is shared whereas applicant-nonapplicant participants only share the source. Future research using other-ratings or objective measures of workplace criteria would be able to address this limitation. Third, there were some baseline differences between the groups on gender and age and there may be other unmeasured differences. Nonetheless, the sampling procedure meant that both groups tended to come from a broad cross-section of society. Furthermore, the size of the differences between applicants and the minimal effect that controlling for covariates had on these effects, suggests that almost all of the observed differences can be attributed to the motivational demands of the applicant context. Fourth, using two different, albeit highly correlated, measures of the HEXACO at baseline and follow-up places some limitations on the kinds of analyses of residuals and faking that can be performed. Finally, larger sample sizes would have been beneficial for estimating tighter confidence intervals on any difference in predictive validity.

4.4. Conclusion

Using a large sample, unique design, and methodological refinements, the current study contributes to an understanding of the effect of applicant response distortion on predictive validity of personality in applicant settings. In particular, the study overcomes limitations by using a predictive/time-lagged design for both applicants and non-applicants, and the inclusion of a follow-up measure of personality is a strength. It also contributes to understanding of the ability of HEXACO personality to predict beneficial and "dark side" workplace behaviors.

In particular, the study provides useful information for practitioners and organizations considering using the HEXACO-PI-R model of personality in selection settings.

5. Open Practices

Data, data analysis scripts, supplementary materials, and item-level information are provided at https://osf.io/wa6yj. The study was not preregistered.

References

Anglim, J., Bozic, S., Little, J., & Lievens, F. (2018). Response distortion on personality tests in applicants: comparing high-stakes to low-stakes medical settings. *Advances in Health Sciences Education, 23*, 311-321.

Anglim, J., & Grant, S. L. (2014). Incremental criterion prediction of personality facets over factors: Obtaining unbiased estimates and confidence intervals. *Journal of Research in Personality, 53*, 148-157.

Anglim, J., & Grant, S. L. (2016). Predicting psychological and subjective well-being from personality: Incremental prediction from 30 facets over the big 5. *Journal of Happiness studies, 17*, 59-80.

Anglim, J., Morse, G., De Vries, R. E., MacCann, C., & Marty, A. (2017). Comparing Job Applicants to Non-applicants Using an Item-level Bifactor Model on the HEXACO Personality Inventory. *European Journal of Personality, 31*, 669-684.

Anglim, J., & O'Connor, P. (2018). Measurement and research using the Big Five, HEXACO, and narrow traits: A primer for researchers and practitioners. *Australian Journal of Psychology*.

Ashton, M. C. (1998). Personality and job performance: The importance of narrow traits. *Journal of Organizational Behavior, 19*(3), 289-303.

Ashton, M. C., Lee, K., & de Vries, R. E. (2014). The HEXACO Honesty-Humility, Agreeableness, and Emotionality Factors A Review of Research and Theory. *Personality and Social Psychology Review, 18*(2), 139-152.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta‐analysis. *Personnel Psychology, 44*(1), 1-26.

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology, 81*(3), 261-272.

Barrick, M. R., & Mount, M. K. (2012). Nature and use of personality in selection. In N. Schmitt (Ed.), *The Oxford Handbook of Personnel Assessment and Selection.* (pp. 225-251). New York, NY, US: Oxford University Press.

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

Bennett, R. J., & Robinson, S. L. (2000). Development of a measure of workplace deviance. *Journal of Applied Psychology, 85*(3), 349-360.

Berry, C. M., Carpenter, N. C., & Barratt, C. L. (2012). Do other-reports of counterproductive work behavior provide an incremental contribution over self-reports? A meta-analytic comparison. *Journal of Applied Psychology, 97*(3), 613-636.

Berry, C. M., Ones, D. S., & Sackett, P. R. (2007). Interpersonal deviance, organizational deviance, and their common correlates: a review and meta-analysis. *Journal of Applied Psychology, 92*(2), 410-424.

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*(4), 317-335.

Campbell, J. P. (1990). An overview of the army selection and classification project (Project A). *Personnel Psychology, 43*(2), 231-239.

Ceschi, A., Sartori, R., Dickert, S., & Costantini, A. (2016). Grit or honesty-humility? New insights into the moderating role of personality between the health impairment process and counterproductive work behavior. *Frontiers in psychology, 7*, 1799.

Chiaburu, D. S., Oh, I.-S., Berry, C. M., Li, N., & Gardner, R. G. (2011). The five-factor model of personality traits and organizational citizenship behaviors: a meta-analysis. *Journal of Applied Psychology, 96*(6), 1140-1166.

Chirumbolo, A. (2015). The impact of job insecurity on counterproductive work behaviors: The moderating role of honesty–humility personality trait. *The Journal of psychology, 149*(6), 554-569.

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*(4), 847-860.

Christiansen, N. D., & Robie, C. (2011). Further consideration of the use of narrow trait scales. *Canadian Journal of Behavioural Science, 43*(3), 183-194.

Conte, J. M., & Gintoft, J. N. (2005). Polychronicity, Big Five personality dimensions, and sales performance. *Human Performance, 18*(4), 427-444.

Costa, P. T., & MacCrae, R. R. (1992). *Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO FFI): Professional manual*: Psychological Assessment Resources.

Crant, J. M. (1995). The Proactive Personality Scale and objective job performance among real estate agents. *Journal of Applied Psychology, 80*(4), 532-537.

Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions*. Oxford, England: University of Illinois Press.

de Vries, A., de Vries, R. E., & Born, M. P. (2011). Broad versus narrow traits: Conscientiousness and Honesty–Humility as predictors of academic criteria. *European Journal of Personality, 25*(5), 336-348.

de Vries, R. E., & van Gelder, J.-L. (2015). Explaining workplace delinquency: The role of Honesty–Humility, ethical culture, and employee surveillance. *Personality and Individual Differences, 86*, 112-116.

De Vries, R. E., Zettler, I., & Hilbig, B. E. (2014). Rethinking trait conceptions of social desirability scales: Impression management as an expression of honesty-humility. *Assessment, 21*(3), 286-299.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93*(5), 880-896.

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73*(6), 1246-1256.

Douglas, E. F., McDaniel, M. A., & Snell, A. F. (1996). *The validity of non-cognitive measures decays when applicants fake.* Paper presented at the Academy of Management Proceedings.

Ellingson, J. E., Sackett, P. R., & Connelly, B. S. (2007). Personality assessment across selection and development contexts: Insights into response distortion. *Journal of Applied Psychology, 92*(2), 386-395.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American psychologist, 48*(1), 26-34.

Grieve, R., & De Groot, H. T. (2011). Does online psychological test administration facilitate faking? *Computers in Human Behavior, 27*(6), 2386-2391.

Hough, L. M. (1997). Personality at work: Issue and evidence. In M. D. Hakel (Ed.), *Beyond multiple choice: Evaluating alternatives to traditional testing for selection* (pp. 131-166). Mahwah, NJ: Lawrence Erlbaum Associates.

Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology, 75*(5), 581-595.

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: the Big Five revisited. *Journal of Applied Psychology, 85*(6), 869-879.

Jenkins, M., & Griffith, R. (2004). Using personality constructs to predict performance: Narrow or broad bandwidth. *Journal of Business & Psychology, 19*(2), 255-269.

Jeong, Y. R., Christiansen, N. D., Robie, C., Kung, M. C., & Kinney, T. B. (2017). Comparing applicants and incumbents: Effects of response distortion on mean scores and validity of personality measures. *International Journal of Selection and Assessment, 25*(3), 311-315.

Judge, T. A., Rodell, J. B., Klinger, R. L., Simon, L. S., & Crawford, E. R. (2013). Hierarchical representations of the five-factor model of personality in predicting job performance: integrating three organizing frameworks with two theoretical perspectives. *Journal of Applied Psychology, 98*(6), 875-925.

Komar, S., Brown, D. J., Komar, J. A., & Robie, C. (2008). Faking and the validity of conscientiousness: A Monte Carlo investigation. *Journal of Applied Psychology, 93*(1), 140-154.

Lee, K., & Allen, N. J. (2002). Organizational citizenship behavior and workplace deviance: the role of affect and cognitions. *Journal of Applied Psychology, 87*(1), 131-142.

Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO personality inventory. *Multivariate Behavioral Research, 39*(2), 329-358.

Lee, K., & Ashton, M. C. (2008). The HEXACO personality factors in the indigenous personality lexicons of English and 11 other languages. *Journal of personality, 76*(5), 1001-1054.

Lee, K., Ashton, M. C., & de Vries, R. E. (2005). Predicting workplace delinquency and integrity with the HEXACO and five-factor models of personality structure. *Human Performance, 18*(2), 179-197.

Lee, K., Ashton, M. C., Morrison, D. L., Cordery, J., & Dunlop, P. D. (2008). Predicting integrity with the HEXACO personality model: Use of self‐ and observer reports. *Journal of Occupational and Organizational Psychology, 81*(1), 147-167.

Lounsbury, J. W., Gibson, L. W., & Hamrick, F. L. (2004). The development and validation of a personological measure of work drive. *Journal of Business and Psychology, 18*(4), 427-451.

MacCann, C. (2013). Instructed faking of the HEXACO reduces facet reliability and involves more Gc than Gf. *Personality and Individual Differences, 55*(7), 828-833.

Marcus, B., Lee, K., & Ashton, M. C. (2007). Personality dimensions explaining relationships between integrity tests and counterproductive behavior: Big Five, or one in addition? *Personnel Psychology, 60*(1), 1-34.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007a). Are we getting fooled again? Coming to terms with limitations in the use of personality tests for personnel selection. *Personnel Psychology, 60*(4), 1029-1049.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007b). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology, 60*(3), 683-729.

Mueller-Hanson, R., Heggestad, E. D., & Thornton III, G. C. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. *Journal of Applied Psychology, 88*(2), 348-355.

Musek, J. (2007). A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality, 41*(6), 1213-1233.

Oh, I. S., Lee, K., Ashton, M. C., & De Vries, R. E. (2011). Are Dishonest Extraverts More Harmful than Dishonest Introverts? The Interaction Effects of Honesty‐Humility and Extraversion in Predicting Workplace Deviance. *Applied Psychology, 60*(3), 496-516.

Ones, D. S., & Viswesvaran, C. (1996). Bandwidth–fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior, 17*(6), 609-626.

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*(6), 660-679.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology, 78*(4), 679-703.

Paunonen, S. V., & Ashton, M. C. (2001). Big five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology, 81*(3), 524-539.

Paunonen, S. V., & Jackson, D. N. (2000). What is beyond the big five? Plenty! *Journal of personality, 68*(5), 821-835.

Peterson, M. H., Griffith, R. L., Isaacson, J. A., O'Connell, M. S., & Mangos, P. M. (2011). Applicant faking, social desirability, and the prediction of counterproductive work behaviors. *Human Performance, 24*(3), 270-290.

Revelle, W. (2018). psych: Procedures for Personality and Psychological Research. Northwestern University, Evanston, Illinois, USA. Retrieved from https://cran.r-project.org/package=psych

Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology, 58*(1), 103-139.

Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support? *Human resource management review, 16*(2), 155-180.

Salgado, J. (1997). The Five Factor Model of personality and job performance in the European Community. *Journal of Applied Psychology, 82*(1), 30-43.

Salgado, J. F., Moscoso, S., & Berges, A. (2013). Conscientiousness, its facets, and the prediction of job performance ratings: Evidence against the narrow measures. *International Journal of Selection and Assessment, 21*(1), 74-84.

Salgado, J. F., Moscoso, S., Sanchez, J. I., Alonso, P., Choragwicka, B., & Berges, A. (2015). Validity of the five-factor model and their facets: The impact of performance measure and facet residualization on the bandwidth-fidelity dilemma. *European Journal of Work and Organizational Psychology, 24*(3), 325-349.

Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Metaanalyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology, 37*(3), 407-422.

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*(3), 613.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: a meta‐analytic review. *Personnel Psychology, 44*(4), 703-742.

Tett, R. P., Steele, J. R., & Beauregard, R. S. (2003). Broad and narrow measures on both sides of the personality–job performance relationship. *Journal of Organizational Behavior, 24*(3), 335-356.

Uziel, L. (2010). Rethinking social desirability scales from impression management to interpersonally oriented self-control. *Perspectives on Psychological Science, 5*(3), 243-262.

Van der Linden, D., te Nijenhuis, J., & Bakker, A. B. (2010). The general factor of personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of Research in Personality, 44*(3), 315-327.

Vinchur, A. J., Schippmann, J. S., Switzer III, F. S., & Roth, P. L. (1998). A meta-analytic review of predictors of job performance for salespeople. *Journal of Applied Psychology, 83*(4), 586-597.

Ziegler, M., Bensch, D., Maaß, U., Schult, V., Vogel, M., & Bühner, M. (2014). Big Five facets as predictor of job training performance: The role of specific job demands. *Learning and Individual Differences, 29*, 1-7.

Table 1

Reliability, Descriptive Statistics and Intercorrelations Demographics, HEXACO Personality, OCB and CWB for Applicants (upper-diagonal) and Non-Applicants (lower-diagonal)

	1	2	3	4	5	6	7	8	9	10	11
1. CWB Total		-.22	-.32	-.02	-.09	-.23	-.22	-.04	.18	-.24	.01
2. OCB Total	-.28		.14	-.15	.34	.31	.26	.07	-.11	-.08	.05
3. Honesty-Humility	-.25	.17		-.16	.19	.41	.31	.30	-.05	.24	-.05
4. Emotionality	.07	-.04	-.02		-.16	-.23	-.05	-.12	-.24	-.06	.10
5. Extraversion	-.30	.39	.15	-.15		.47	.51	.18	-.07	-.21	.06
6. Agreeableness	-.28	.27	.41	-.19	.35		.45	.21	.03	-.08	-.06
7. Conscientiousness	-.33	.31	.22	-.09	.29	.17		.18	-.15	-.34	-.07
8. Openness	-.06	.13	.02	-.08	.28	.09	.02		.08	.05	.14
9. Male	.00	-.01	-.13	-.29	-.02	.10	.02	-.05		.17	-.07
10. Age	-.26	.15	.17	-.20	.25	.13	.14	.08	.29		-.09
11. Bachelor or higher	.13	-.03	-.09	-.03	.03	-.08	-.11	.19	.04	-.15	
Alpha Non-Applicant	.85	.90	.90	.87	.93	.90	.88	.88			
Alpha Applicant	.86	.83	.88	.82	.91	.89	.88	.89			
Mean Non-Applicant	1.63	5.54	3.73	2.97	3.70	3.16	3.69	3.62	0.60	50.71	0.78
SD Non-Applicant	0.59	0.87	0.48	0.44	0.52	0.47	0.43	0.46	0.49	10.35	0.41
Mean Applicant	1.62	5.63	4.06	2.98	4.03	3.61	3.97	3.64	0.31	41.88	0.68
SD Applicant	0.60	0.67	0.36	0.34	0.37	0.39	0.35	0.43	0.46	12.14	0.47
d	-0.01	0.11	0.68	0.04	0.64	0.94	0.65	0.04	-0.60	-0.85	-0.25
d lower 95% CI	-0.22	-0.10	0.46	-0.18	0.42	0.72	0.44	-0.17	-0.82	-1.07	-0.46
d Upper 95% CI	0.20	0.32	0.90	0.25	0.85	1.17	0.87	0.25	-0.39	-0.63	-0.04
sig			*		*	*	*		*	*	*

Note. Male is coded 0 = Female; 1 = Male. Bachelor or higher educational qualification is coded yes = 1, no = 0. Personality is on a 1 to 5 scale; OCB and CWB is on a 1 to 7 scale. Significance test for group differences is based on two-tailed between-subjects t-test. Significant correlations ($p < .05$) are those greater than or equal to .13 for the applicant sample and .11 for the non-applicant sample. The sample size for descriptive statistics was 607 except for age which had some missing data (n = 514).

* $p < .05$

Table 2

Regression Models Predicting CWB and OCB from Baseline HEXACO personality Domains in Applicants and Non-Applicants

	CWB				OCB			
	b		Beta		b		Beta	
	NonApp	App	NonApp	App	NonApp	App	NonApp	App
Coefficients								
Honesty-humility	-0.14*	-0.45*	-0.11	-0.27	0.05	0.01	0.03	0.00
Emotionality	0.00	-0.14	0.00	-0.08	0.10	-0.14	0.05	-0.07
Extraversion	-0.19*	0.10	-0.17	0.06	0.47*	0.39*	0.28	0.21
Agreeableness	-0.16*	-0.19	-0.13	-0.12	0.25*	0.29*	0.14	0.17
Conscientiousness	-0.33*	-0.21	-0.24	-0.12	0.42*	0.15	0.21	0.08
Openness	0.01	0.09	0.00	0.07	0.08	-0.04	0.04	-0.03
Model Fit								
Standard error	0.53	0.57			0.78	0.62		
Adjusted R^2	.18	.11			.21	.13		

Note. NonApp = Non-Applicant, App = Applicant, OCB = Organizational Citizenship Behavior, CWB = Counterproductive Work Behavior. b is unstandardized coefficient. Beta is standardized coefficient.
* $p < .05$

Table 3

Regression Models Predicting CWB and OCB from Residualized Baseline HEXACO Personality Domains (i.e., Controlling for Follow-Up Personality) in Applicants and Non-Applicants

	CWB				OCB			
	b		Beta		b		Beta	
	NonApp	App	NonApp	App	NonApp	App	NonApp	App
Coefficients								
Resid. Honesty	-0.05	0.04	-0.03	0.02	-0.07	-0.19	-0.03	-0.09
Resid. Emotionality	-0.12	-0.35*	-0.06	-0.16	-0.06	-0.06	-0.02	-0.02
Resid. Extraversion	-0.10	0.07	-0.06	0.04	0.01	-0.18	0.00	-0.08
Resid. Agreeableness	-0.04	-0.05	-0.02	-0.03	0.16	0.31	0.06	0.15
Resid. Conscientiousness	-0.20	-0.02	-0.11	-0.01	0.33*	0.05	0.12	0.02
Resid. Openness	-0.01	-0.05	-0.01	-0.03	-0.01	0.00	0.00	0.00
Model Fit								
Standard error	0.59	0.60			0.87	0.67		
Adjusted R^2	.01	.00			.00	.00		

Note. NonApp = Non-Applicant, App = Applicant, OCB = Organizational Citizenship Behavior, CWB = Counterproductive Work Behavior. Resid. = Residualized. b is unstandardized coefficient. Beta is standardized coefficient.
* $p < .05$

Table 4
Zero-Order and Semi-Partial Correlations between Baseline HEXACO facets and Employee Outcomes for Applicants and Non-Applicants

	Zero-order correlations				Semi-partial correlations			
	Non-Applicants		Applicants		Non-Applicants		Applicants	
	CWB	OCB	CWB	OCB	CWB	OCB	CWB	OCB
H1: Sincerity	**-.23**	.08	**-.26**	.12	-.07	-.07	-.04	-.02
H2: Fairness	**-.31**	**.27**	**-.27**	.17	-.07	.06	-.05	-.03
H3: Greed avoidance	-.10	.06	**-.22**	.07	.09	-.05	.02	.06
H4: Modesty	-.11	.11	**-.22**	.09	.04	.05	.08	-.01
E1: Fearfulness	.03	**-.18**	-.04	-.14	-.04	**-.18**	-.03	.00
E2: Anxiety	**.20**	-.10	.05	-.20	.08	.09	-.01	.01
E3: Dependence	.08	.00	.04	-.10	.02	.03	.04	-.03
E4: Sentimentality	-.14	**.20**	-.12	.09	-.06	.08	.01	.02
X1: Social self-esteem	**-.25**	**.26**	-.14	**.22**	.06	-.14	-.07	-.08
X2: Social Boldness	**-.27**	**.37**	-.05	**.29**	-.10	.14	-.01	.06
X3: Sociability	-.15	**.26**	.00	.20	.03	-.02	.07	-.08
X4: Liveliness	**-.29**	**.36**	-.11	**.37**	.02	.01	.00	.10
A1: Forgiveness	**-.19**	**.26**	**-.21**	**.30**	.01	.09	-.08	.07
A2: Gentleness	**-.22**	.14	-.18	**.28**	-.03	-.08	.01	.10
A3: Flexibility	**-.23**	**.19**	-.18	.16	-.06	-.01	.01	-.11
A4: Patience	**-.22**	**.25**	-.15	**.23**	.08	-.02	.07	-.08
C1: Organization	**-.24**	.17	-.18	**.22**	-.01	-.06	-.02	.04
C2: Diligence	**-.35**	**.44**	-.17	**.30**	-.12	**.25**	-.09	.09
C3: Perfectionism	-.09	.12	-.03	.08	.11	-.04	.18	-.11
C4: Prudence	**-.28**	.17	**-.26**	.20	.02	-.13	-.09	-.02
O1: Aesthetic appreciation	-.10	.03	-.16	.05	-.07	-.15	-.16	-.02
O2: Inquisitiveness	-.01	.14	-.02	.06	.04	.08	.02	-.02
O3: Creativity	-.09	.15	-.07	.10	-.04	.03	-.05	.03
O4: Unconventionality	.03	.08	.15	.01	.06	.03	**.22**	.00
I: Altruism	**-.25**	**.32**	-.20	**.27**	-.09	**.17**	.02	.11

Note. OCB = Organizational Citizenship Behavior, CWB = Counterproductive Work Behavior. Semi-partial Correlations are those between Baseline HEXACO facets and Employee Outcomes Controlling for the Six HEXACO Domains. Correlations bolded are statistically significant (p < .001).

6.   Online Supplement

6.1. Robustness checks that group differences in HEXACO personality are not due to age or gender

To examine the degree to which the above differences were caused by the applicant context and not underlying group differences, we performed two additional analyses. First, a regression was run for each domain score and each facet score as an outcome measure predicted by context (applicant versus non-applicant) and the demographic factors of gender, age and age-squared. The Cohen's d effect sizes for context in these models were quite similar to those presented in Table 1 without covariates. The average increase in absolute effect size when covariates were included was .09 for domains and .08 for facets (i.e., absolute covariate adjusted Cohen's d minus absolute Cohen's d without adjustment). The average increase in effect sizes when covariates were included was close to zero (i.e., .0004 for domains and .002 for facets; covariate adjusted Cohen's d minus Cohen's d without adjustment). The main noteworthy change was that the emotionality effect increased by .20 when covariates were included, presumably caused by the greater proportion of women in the job applicant sample and that women tend to score higher on emotionality. Second, we examined differences between applicants and non-applicants on the personality domain scores of the SACS 6, which was administered at time 2 in a pure research setting. Minimal differences were observed between the two groups with an average Cohen's d of .14 across the six domains. Applicants scored slightly higher on SACS-6 emotionality (d = .28) and agreeableness (d = .36). Nonetheless, overall, the results suggest that the vast majority of the estimated effects were due to actual context effects.

6.2. Robustness checks that differences in predictive validity are not influenced by time between testing sessions

We also examined whether the time between time 1 and time 2 testing influenced the results. We fit regression models predicting each time 2 domain personality score from (a) the time 1 domain personality score, (b) the log of days between time 1 and time 2 testing, and (c) the interaction of these two variables. If personality changed progressively over time, then we would expect the time by time 1 personality interaction effect to be negative reflecting the decreased stability of personality over longer periods of time. The results showed no significant effect for either time or the time by time1 personality interaction. Thus, the stability of personality did not significantly vary as a function of the test intervals studied in the present study.

6.3. Differences between applicant and non-applicant HEXACO scale variances

In addition to the effect of applicant context on personality test means, standard deviations tended to be smaller in the job applicant context. Standard deviations for applicants were on average only 80% for domains and 79% for facets compared to non-applicant context. Based on Levene's test all domains except for openness had significantly different standard deviations, and all but fearfulness, inquisitiveness, and creativity were significantly different at the facet level. The correlation over scales of the absolute Cohen's d effect size with the ratio of applicant to non-applicant standard deviations was moderate and negative ($r = -.43$). This is consistent with changes in mean leading to mild floor and ceiling effects. However, given the relatively weak correlation, it is also likely that for some items the mid-point is perceived as more socially desirable than the extremes.

6.4. Differences between applicant and non-applicant HEXACO factor loadings

Several analyses were performed to verify that the factor structure of the HEXACO was consistent with theory and to examine differences between job applicants and non-applicants. In both the non-applicant and applicant samples, parallel analysis suggested that six factors were appropriate for representing the 25 facets. We performed maximum likelihood factor analysis with oblique promax rotation extracting six factors. In the non-applicant sample, all 24 aligned facets loaded maximally on theorized factors with only three scales cross-loading above .30. In the applicant sample, 22 of 24 items loaded maximally on their theorized factor, and only four scales cross-loaded above .30. Thus, overall, there was good convergence with the theorized structure.

6.5. Differences between applicant and non-applicant in evidence for global evaluative factor

Consistent with a more influential general evaluative factor driving applicant responses, the global factor of personality explained more variance in job applicants than in non-applicants. When principal component analysis was performed both on the 25 facets and on the six domains, the variance explained by the first component was greater for job applicants (facets = 26.7%, domains = 39.9%) than for non-applicants (facets = 21.0%, domains = 32.0%). Similarly, the average absolute correlation between facets and between factors was larger for the applicants (facets = .23, domains = .26) than for the non-applicants (facets = .18, domains = .17).

Table S1

Reliability, Descriptive Statistics for Baseline HEXACO Personality Facets for Applicants and

Non-Applicants

Variable	Non-Applicants			Applicants				
	Alpha	M	SD	Alpha	M	SD	d	Sig
H1: Sincerity	.78	3.65	0.58	.75	3.91	0.46	0.45	*
H2: Fairness	.83	4.13	0.66	.73	4.52	0.41	0.59	*
H3: Greed avoidance	.85	3.50	0.70	.82	3.80	0.61	0.43	*
H4: Modesty	.81	3.66	0.61	.71	4.01	0.43	0.59	*
E1: Fearfulness	.77	2.62	0.64	.73	2.62	0.56	-0.01	
E2: Anxiety	.84	3.00	0.72	.80	2.76	0.59	-0.33	*
E3: Dependence	.78	2.79	0.59	.69	2.92	0.49	0.21	*
E4: sentimentality	.77	3.46	0.59	.67	3.64	0.44	0.31	*
X1: Social self-esteem	.85	4.08	0.58	.82	4.42	0.39	0.59	*
X2: Social Boldness	.82	3.65	0.65	.80	3.83	0.51	0.27	*
X3: Sociability	.88	3.32	0.75	.81	3.78	0.53	0.60	*
X4: Liveliness	.84	3.75	0.64	.79	4.10	0.44	0.56	*
A1: Forgiveness	.86	2.93	0.69	.82	3.51	0.54	0.85	*
A2: Gentleness	.75	3.10	0.57	.73	3.50	0.49	0.71	*
A3: Flexibility	.66	3.18	0.52	.65	3.60	0.46	0.81	*
A4: Patience	.84	3.44	0.66	.77	3.81	0.50	0.57	*
C1: Organization	.85	3.67	0.72	.85	4.02	0.56	0.48	*

C2: Diligence	.78	3.84	0.55	.76	4.06	0.42	0.40*
C3: Perfectionism	.78	3.58	0.59	.75	3.79	0.50	0.36*
C4: Prudence	.80	3.66	0.56	.77	4.01	0.42	0.63*
O1: Aesthetic appreciation	.82	3.59	0.71	.79	3.72	0.60	0.19*
O2: Inquisitiveness	.79	3.85	0.63	.79	3.86	0.58	0.01
O3: Creativity	.77	3.51	0.62	.77	3.57	0.56	0.10
O4: Unconventionality	.75	3.53	0.55	.71	3.39	0.45	-0.24*
I: Altruism	.79	3.98	0.53	.78	4.26	0.41	0.53*

Note. All facets are on a 1 to 5 scale. Significance test is based on between-subjects t-test.

* $p < .05$

Table S2

Regression Models Predicting CWB and OCB from Baseline HEXACO personality Domains and

Demographics in Applicants and Non-Applicants

	CWB				OCB			
	Coef		Beta		Coef		Beta	
	NonApp	App	NonApp	App	NonApp	App	NonApp	App
Coefficients								
Honesty	-0.11	-0.31*	-0.09	-0.19	0.04	0.01	0.02	0.01
Emotionality	-0.02	-0.05	-0.01	-0.03	0.11	-0.19	0.05	-0.10
Extraversion	-0.16*	0.09	-0.14	0.05	0.46*	0.37	0.27	0.20
Agreeableness	-0.17*	-0.21	-0.14	-0.13	0.26*	0.31*	0.14	0.18
Conscientiousness	-0.32*	-0.30*	-0.24	-0.18	0.42*	0.10	0.21	0.05
Openness	0.01	0.08	0.00	0.06	0.08	-0.02	0.04	-0.01
Male	0.04	0.22*	0.03	0.17	-0.02	-0.17	-0.01	-0.12
Age	-0.007*	-0.012*	-0.13	-0.25	-0.003	-0.002	0.03	-0.04
Model Fit								
Standard error	.53	.54			.78	.62		
Adjusted R2	.19	.18			.20	.14		

Note. NonApp = Non-Applicant, App = Applicant, OCB = Organizational Citizenship Behavior,

CWB = Counterproductive Work Behavior. b is unstandardized coefficient. Beta is standardized

coefficient. Male is coded 0 = female, male = 1.

* $p < .05$

Table S3

Test of differences between applicants and non-applicants for correlations of personality with outcomes

Outcome	Trait	Correlation (r)				95% CI	
		Non-Applicants	Applicants	Difference	p	Lower	Upper
CWB	Honesty-humility	-.25	-.32	-.07	.34	-.08	.22
OCB	Honesty-humility	.17	.14	-.03	.67	-.12	.19
CWB	Emotionality	.07	-.02	-.09	.27	-.07	.25
OCB	Emotionality	-.04	-.15	-.11	.18	-.05	.27
CWB	Extraversion	-.30	-.09	.21	.01	-.36	-.06
OCB	Extraversion	.39	.34	-.05	.45	-.09	.20
CWB	Agreeableness	-.28	-.23	.05	.54	-.20	.10
OCB	Agreeableness	.27	.31	.04	.59	-.19	.11
CWB	Conscientiousness	-.33	-.22	.12	.12	-.27	.03
OCB	Conscientiousness	.31	.26	-.05	.52	-.10	.20
CWB	Openness	-.06	-.04	.02	.81	-.18	.14
OCB	Openness	.13	.07	-.06	.44	-.10	.22

Note. OCB = Organizational citizenship behavior; CWB = Counterproductive work behavior.