

ABSTRACT

YOUNG, AMANDA LEIGH. Faking Resistance of a Forced-Choice Measure of the Dark Triad. (Under the direction of Dr. Adam W. Meade).

Personality assessments are used by organizations for important administrative purposes, such as hiring and promotion. Research has primarily been focused on bright side personality traits, but the measurement of dark traits has the potential to offer information regarding an important part of the performance domain: counterproductive work behaviors. Dark personality traits are those that reflect a tendency to serve one's own goals while harming others. The most commonly studied dark traits in organizational sciences and the focus of the present study are Machiavellianism, narcissism, and psychopathy, collectively referred to as the Dark Triad (DT). Despite the potential benefits of dark personality assessments, concerns about the accuracy of measurement that have surrounded bright side personality assessments are potentially even more problematic for dark traits. These traits are not socially desirable, and standard survey methods allow people to purposefully deflate their scores on dark traits in order to present themselves more positively. This is known as "faking good." A potential solution to faking good is the use of the forced-choice method to design assessments that are more difficult to fake. This study administered new multidimensional forced-choice (MFC) and single-stimulus measures of the DT. Participants were able to alter their scores when asked to respond like an ideal job applicant, and counter to expectations, the MFC format did not decrease faking. Thus, the forced-choice method may not always be an effective way to reduce faking. The results showed that the DT, especially psychopathy, explained additional variance in counterproductive work behaviors beyond bright side personality traits. This reinforces the importance of measuring dark traits, so future research should continue to explore ways to more accurately measure dark personality.

© Copyright 2018 by Amanda Leigh Young

All Rights Reserved

Faking Resistance of a Forced-Choice Measure of the Dark Triad

by
Amanda Leigh Young

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Psychology

Raleigh, North Carolina
2018

APPROVED BY:

Dr. Adam W. Meade
Committee Chair

Dr. Samuel B. Pond, III

Dr. S. Bartholomew Craig

Dr. Mark A. Wilson

BIOGRAPHY

Amanda Young was born in the small town of Newton, North Carolina to parents George and Mary Beth Young. She attended a high school in the country that had its own farm. In high school she developed a passion for tennis and music performance. Not knowing what she wanted to do with her life, she went on to complete her undergraduate education at Lenoir-Rhyne University and was a psychology major and a music minor. While at Lenoir-Rhyne, she played flute in several musical ensembles and was on the tennis team. She graduated with a B.A. in Psychology on a very windy day in May 2014. Upon enrolling in the Industrial/Organizational Psychology doctoral program at North Carolina State University Amanda forgot what having hobbies is like, but she completed her M.S. in 2016 and Ph.D. in 2018.

ACKNOWLEDGMENTS

First, I would like to thank my committee members: Dr. Adam Meade, Dr. Bart Craig, Dr. Mark Wilson, and Dr. Bob Pond for their support and feedback throughout this process. A special thanks to my committee chair, Adam, for helping me move from several less than spectacular ideas to a study that I wouldn't be embarrassed to tell people about. Additionally, whether he remembers it or not, I would like to thank Bart for always being willing to meet with me and discuss big ideas and offer practical guidance about cars or my dissertation.

I would also like to thank all of the other students at NCSU (and Peter) who kindly agreed to help with the validation of my scale. Special thanks to everyone who gave me feedback on my many less than perfect drafts: Nicole Francavilla, Emery Clayton, Catherine Neale, and Unber Ahmad. An extra special thanks to my friend Unber not only for her support but for always pushing me to do better, even though I don't always listen.

Thank you to my parents, brother and the rest of my family who have not only dealt with me through the process but have been supportive. I would also like to express my gratitude for all of the friends who have added some enjoyment to life as a graduate student. I am grateful to have people who still like me even after they had to informally complete and give me feedback on a personality assessment... several times. Finally, thank you to my fiancé, Chris, who continued to ground and support me, even when I was at my craziest.

TABLE OF CONTENTS

| | |
|---|----|
| LIST OF TABLES | v |
| Faking Resistance of a Forced-Choice Measure of the Dark Triad..... | 1 |
| Dispositional Antecedents of CWBs | 2 |
| HEXACO Model | 4 |
| Dark Personality..... | 5 |
| Facets of DT traits..... | 6 |
| Personality Measurement..... | 8 |
| Models of Faking | 10 |
| Faking and dark personality..... | 11 |
| Reducing Faking..... | 12 |
| MFC Scoring..... | 14 |
| Choosing a model | 15 |
| Current Paper | 16 |
| Method | 18 |
| Participants..... | 19 |
| Procedure | 19 |
| Measures | 20 |
| Dark Personality..... | 21 |
| Bright Personality | 21 |
| CWBs..... | 22 |
| Results..... | 22 |
| Dark Personality Scoring | 22 |
| Test of Hypotheses..... | 24 |
| Research Question | 26 |
| Discussion..... | 27 |
| Limitations and Future Directions | 33 |
| Conclusions..... | 34 |
| References..... | 35 |
| Appendix: Scale Development | 68 |

LIST OF TABLES

| | | |
|----------|--|----|
| Table 1 | <i>Item pairs</i> | 49 |
| Table 2 | <i>Fit Indices for CFAs</i> | 52 |
| Table 3 | <i>Narcissism fit for the Graded Response Model</i> | 53 |
| Table 4 | <i>Psychopathy model fit for the Graded Response Model</i> | 54 |
| Table 5 | <i>Machiavellianism model fit for the Graded Response Models</i> | 55 |
| Table 6 | <i>Means, standard deviations, and correlations of faking Dark Triad scores</i> | 56 |
| Table 7 | <i>Means, standard deviations, and correlations of honest Dark Triad scores</i> | 57 |
| Table 8 | <i>Means, standard deviations, and correlations of HEXACO and CWB scores</i> | 58 |
| Table 9 | <i>Comparison of faking and honest scores</i> | 59 |
| Table 10 | <i>Hypothesis 1 results</i> | 60 |
| Table 11 | <i>Hypothesis 2 results</i> | 61 |
| Table 12 | <i>Hypothesis 3 results</i> | 62 |
| Table 13 | <i>Research Question results</i> | 63 |
| Table 14 | <i>Frequency of choosing the narcissism option in the MFC assessment in the faking condition</i> | 64 |
| Table 15 | <i>Levene's test for equality of variance</i> | 65 |
| Table 16 | <i>Regression slopes for HEXACO predicting faking scores</i> | 66 |

Faking Resistance of a Forced-Choice Measure of the Dark Triad

In recent years there has been growing interest in the negative workplace behaviors brought to light by highly publicized scandals (Wu & LeBreton, 2011). These events have spurred organizational scholars to examine the antecedents of these counterproductive work behaviors (CWBs; Spain, Harms, & LeBreton, 2014). Many researchers have argued that the best way to predict negative behaviors is to focus on negative, or “dark,” personality traits, such as narcissism, psychopathy, and Machiavellianism (Wu & LeBreton, 2011), which have been shown to have a significant relationship with CWBs (O’Boyle, Forsyth, Banks, & McDaniel 2012).

However, the usefulness of assessing dark personality traits in organizational settings is limited by problems with accurate measurement (Guenole, 2014). Items reflecting dark personality traits are clearly not socially desirable, which makes it easy for individuals to “fake good” and receive lower scores than their true scores on these scales. This makes the use of traditional self-report measures of dark personality traits in organizations suspect, especially in high stakes selection testing when applicants are motivated to present themselves positively.

A promising solution to this issue is designing the assessments harder to fake by incorporating a forced-choice design (Guenole, 2014). Assessments using a forced-choice design present multiple items simultaneously and require respondents to choose the item that is most like them amongst the given options. This design makes it more difficult for respondents to simply report all high (or low in the case of negative traits) ratings and inflate (or deflate) their scores. These designs have not been widely implemented in the past because of the psychometric problems associated with ipsative scoring, but new advances in Item Response Theory models have overcome these problems (Brown & Maydeu-Olivares, 2012; Stark, Chernyshenko, &

Drasgow, 2005). However, this method has not yet been implemented for the measurement of important dark personality traits. Thus, the purpose of the present paper is to (1) determine whether a the forced-choice method reduces the ability of individuals to fake good and (2) determine whether dark traits can provide incremental validity in predicting CWBs beyond bright traits.

Dispositional Antecedents of CWBs

Counterproductive work behaviors (CWBs) are “volitional acts that harm or intend to harm organizations and their stakeholders” (Spector & Fox, 2005, p. 151). Organizational theorists adopt an interactionist perspective to understand the antecedents of CWBs. Researchers recognize the importance of both environmental influences, individual difference variables, and their interaction in predicting CWBs (e.g., Sackett & DeVore, 2001). The environment is the context within which individual differences are expressed. Strong situations with rigid roles constrain the expression of individual differences, whereas individual differences have more influence on behaviors in weak situations (Mischel, 1986). Several environmental characteristics have been shown to relate to CWBs including organizational culture, leadership style, and opportunity to engage in counterproductive behaviors (e.g., Greenberg, 2002; Martinko, Gundlach, & Douglas, 2002).

Although both individual difference and environmental variables are necessary for a complete understanding of the antecedents of CWBs, this study will focus on individual difference variables. A focus on individual difference variables is warranted because they have direct implications for selection. Individual differences can be assessed during the selection process, and organizations can use these assessments to select candidates who are less likely to

engage in CWBs (Spain et al., 2014). Further, without an accurate understanding of individual difference variables, it is impossible to determine how they interact with situational factors.

Much of the existing research relating individual differences to CWBs has used the Five Factor Model (FFM), which is a bright side model of personality (Wu & LeBreton, 2011). This model has a hierarchical structure and is comprised of five broad domains: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism; each domain contains several facets (Costa & McCrae, 1995). Neuroticism is defined as a tendency to experience distress (McCrae & John, 1992). Extraversion and agreeableness are both understood as dimensions of interpersonal behavior. Extraversion is a tendency to seek social stimulation, and agreeableness is a tendency to be compassionate rather than antagonistic (Costa, McCrae, & Dye, 1991). Finally, conscientiousness is defined as the combination of self-discipline and need for achievement (Costa et al., 1991). Researchers have found that three of these five domains have a consistent, significant relationship with CWBs. Agreeableness and conscientiousness tend to negatively correlate with CWBs, and neuroticism tends to positively correlate with CWBs (Berry, Ones, & Sackett, 2007; Sackett, Berry, Wiemann, & Laczko, 2006). This information has been indirectly used for employee selection, as integrity tests used in selection tend to be saturated with aspects of conscientiousness, agreeableness, and neuroticism (Wu & LeBreton, 2011). However, several authors have argued that the FFM is a deficient personality framework, especially for the study of the antecedents of CWBs (Spain et al., 2014; Wu & LeBreton, 2011).

The way the FFM was developed specifically excluded dark traits, which are dysfunctional characteristics that represent a middle ground between normal personality and clinical disorder (Spain et al., 2014). It was developed using a lexical approach, which assumes that important personality traits are encoded in language, so the lexicon of a language should

contain words that represent the full range of relevant personality traits (Lee, Ashton, & de Vries, 2005). Researchers utilizing this approach specifically ignored negative or evaluative words, such as evil or dangerous, which precluded the inclusion of any dark traits (Spain et al., 2014). Although many researchers agree that the FFM is not comprehensive (e.g., Lee & Ashton, 2014; Spain et al., 2014), there is not agreement regarding the best way to extend the FFM. Some researchers (e.g., Lee & Ashton, 2014) have addressed the deficiency of the FFM by adding a sixth Honesty–Humility factor (i.e., the HEXACO model) and others by supplementing the FFM with dark personality traits (e.g., Wu & LeBreton, 2011).

HEXACO Model

The HEXACO model is an extension of the FFM that was developed using a similar lexical approach. The HEXACO model extends the FFM because it was based on many languages, whereas the developers of the FFM only used English (Lee & Ashton, 2004). Although there are more subtle differences, the most noticeable difference between the HEXACO model and the earlier five factor models is the addition of the Honesty–Humility dimension. This new dimension is conceptualized as a reluctance versus a willingness to exploit others, which is typically defined by traits such as honesty, fairness, modesty and sincerity (Lee et al., 2005). This factor is especially important for the prediction of CWBs because there is no dimension in the FFM that measures this propensity toward exploiting others. Lee and colleagues (2005) demonstrated the utility of this dimension by showing that the HEXACO model had a stronger relationship with workplace deviance than the FFM, and that the increased predictive power of the HEXACO model was due to the Honesty–Humility dimension.

Although the HEXACO model extends the predictive power of the FFM, researchers have argued that this model is still deficient because of its lack of explicitly dark personality

traits (Wu & LeBreton, 2011). The FFM and HEXACO models are considered *bright side* models because they do not address aberrant traits. Researchers have argued that the omission of dark traits is detrimental because such traits can predict negative behaviors that have both economic and societal implications (Schyns, 2015).

Dark Personality

Dark personality traits are dysfunctional characteristics that are focused on a motivation to enhance the self and harm others (Harms & Spain, 2015). Research regarding dark personality traits has largely focused on a model of dark traits either resembling the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV; APA, 2000) personality disorders or a constellation of three traits (i.e., narcissism, psychopathy, and Machiavellianism) collectively referred to as the Dark Triad (DT; Paulhus & Williams, 2002; Spain et al., 2014). Research on the DT or one of its three component traits has dominated the dark personality literature in the organizational sciences (Harms & Spain, 2015).

The DT is a collection of the similar yet distinct constructs of narcissism, psychopathy, and Machiavellianism. Narcissism and psychopathy were initially studied as subclinical versions of personality disorders (Paulhus & Williams, 2002). A milder level of narcissism (i.e. not warranting a diagnosis of Narcissist Personality Disorder) is considered a normal personality trait that is characterized by vanity and heightened self-admiration (Muris, Merkelbach, Otgaar, & Meijer, 2017). This subclinical level is important because it is related to behaviors pertinent to the workplace, such as exaggerating achievements, rejecting criticism, and refusing to compromise (O'Boyle et al., 2012).

Psychopathy was based on antisocial personality disorder, but, similar to narcissism, is also considered a personality trait (Levenson, Kiehl, & Fitzpatrick, 1995). People high in

psychopathy are impulsive, lack empathy, and disregard rules (Muris et al., 2017).

Machiavellianism was conceptualized as agreement with statements derived from Niccolò Machiavelli's (1532/1950) *The Prince* (Muris et al., 2017). People who are high in Machiavellianism value the effectiveness of manipulative tactics, value efficiency more than morality, and have a cynical view of human nature (O'Boyle, 2012).

Because narcissism, psychopathy, and Machiavellianism were the most empirically studied aversive personality traits, Paulhus and Williams (2002) combined them under the label "Dark Triad" in order to encourage examining them simultaneously. These traits all share a core of "socially malevolent character with behavior tendencies toward self-promotion, emotional coldness, duplicity, and aggressiveness" (Paulhus & Williams, 2002, p. 557). However, they are also separate traits that can be distinguished by their core features. The core of narcissism is entitlement, the core of psychopathy is callous affect, and the core of Machiavellianism is manipulateness (Wu & LeBreton, 2011). A meta-analysis of 245 samples ($N = 43,907$) by O'Boyle and colleagues (2012) found that all three of the traits in the DT have significant relationships with CWBs. In addition, each trait has been correlated with different, specific types of CWBs. For example, narcissism has been related to CEO unethical behaviors (Amernic & Craig, 2010), psychopathy has been related to diminished levels of corporate responsibility (Boddy, 2010), and Machiavellianism has been related to being perceived as an abusive leader (Jonason, Wee, & Li, 2014).

Facets of DT traits. It is important to note that narcissism, psychopathy, and Machiavellianism are actually multidimensional constructs, despite researchers often discussing them as singular traits in the context of the DT (Wu & LeBreton, 2011). However, there is not universal agreement on the subfacets of each trait. For example, narcissism, as measured by the

Narcissistic Personality Inventory (Raskin & Hall, 1979), has been described as being comprised of anywhere between two and seven facets (Ackerman, Witt, Donnellan, Trzesniewski, & Kashy, 2011). Similarly, psychopathy and Machiavellianism have been described with differing numbers of facets (e.g., Christie & Geis, 1970; Cooke & Michie, 2001; Dahling, Whitaker, & Levy, 2009; Levenson, Kiehl, & Fitzpatrick, 1995). This disagreement regarding the composition of the DT traits is problematic because different facets are likely to be differentially related to organizational outcomes. For example, in a two-factor model of psychopathy, a self-centered impulsivity factor predicted CWBs across contexts, but a fearless dominance predicted beneficial outcomes (Blickle & Schutte, 2017).

In addition, the factor structure of all three of the DT traits measured simultaneously may be different than the factor structure of the traits measured separately. The DT are theorized to be overlapping but distinct traits (Paulhus & Williams, 2002), thus measuring all three would likely yield facets that represent combinations of the traits as well as more distinctive facets.

Examination of items in commonly used scales supports this claim; there are several items across measures of the DT traits that relate to similar aspects of malevolent behavior and are thus likely to intercorrelate strongly (Muris et al., 2017). The similarity of the items has created facets representing the overlap between psychopathy, narcissism, and Machiavellianism that could not be found when examining the facets for each trait separately.

Rogoza and Ciecuch (2018) examined the factor structure of the combination of several common DT measures. They found 12 distinct facets, two of which were composed of positively worded items that were negatively related to dark traits. The facets were: impulsive revengefulness, law of jungle rivalry, ingratiation manipulation, leadership/authority, grandiosity, admiration, foolhardiness, grandiose fantasies, exhibitionism, suspiciousness,

compliance with rules, and grit. Some facets were directly related to hypothesized facets of the three DT traits. For example, they found a facet labeled “leadership/authority,” which several authors have considered as a facet of narcissism (e.g., Ackerman et al., 2011; Emmons, 1984). However, other facets they found reflected overlap in DT constructs. For example, their facet of “manipulation” is common to the Machiavellian facet of amoral manipulation (Dahling, Whitaker, & Levy, 2000) and the psychopathy facet of deceitful interpersonal style (Cooke & Michie, 2001). Collectively, these findings show that the overlap and distinctiveness of DT traits are not currently well-defined, and more research is needed to determine how these traits manifest together.

Personality Measurement

Despite the theoretical relevance of dark personality, especially for predicting CWBs, the use of dark traits in organizations has been hindered by concerns regarding accurate measurement (Spain et al., 2014). Personality is typically measured with self-report assessments (Paulhus & Vazire, 2007). The score on these assessments can be conceptually divided into components representing the true score on the construct of interest and inaccurate self-presentation (Sackett, 2011). Self-presentation resulting from either general or situation-specific erroneous self-perceptions is referred to as self-deception, which is differentiated from conscious impression management (Paulhus & Vazier, 2007).

Self-deception is difficult, if not impossible, to address in the context of self-reported personality. Tett and Simonet (2011) recommend that the best way to address self-deception in a self-report assessment is to identify those types of people who are more likely to self-deceive, such as people high in narcissism, and follow up with more in-depth evaluations. Observer ratings may offer a potential solution for self-deception (e.g., Oh, Wang, & Mount, 2011), but

observer ratings also have practical (e.g., harder to obtain, choosing an appropriate observer) and measurement concerns (Connelly & Ones, 2010; König, Steiner Thommen, Wittwer, & Kleinmann, 2017; Schyns, 2015).

In addition to concerns of unconscious self-presentation, self-report assessments also have problems with deliberate impression management. Faking is a specific type of impression management that refers to the variance due to situationally-specific intentional distortion (Sackett, 2011). It has been defined as “a volitional attempt at increasing one’s score on a personality assessment in order to obtain a desired outcome” (Griffith & Peterson, 2011, p. 291). Altering a score in order to seem like a more desirable job candidate is an example of faking good, which can be differentiated from intentionally trying to look worse, which is referred to as faking bad or malingering (Paulhus & Vazire, 2007). Malingering is a larger problem in the context of clinical diagnosis, so this study will focus specifically on faking good.

Research offers mixed opinions regarding the impact of faking in an organizational context. Some argue that overt faking threatens the validity of inferences based on personality measures, decreases the quality of hiring decisions, and attenuates the relationship between personality and outcomes (e.g., Komar, Brown, Komar, & Robie, 2008; Mueller-Hanson, Heggstad, & Thornton, 2003). For example, Peterson and colleagues (2011) found that removing participants flagged as faking (based on within person differences in faking and honest responding conditions) increased the relationship between conscientiousness and CWBs. However, others argue that faking has little impact on test validity (e.g., Hogan, Barrett, & Hogan, 2007; Ones, Viswesvaran, & Reiss, 1996). For example, Watts and colleagues (2016) found little evidence that the relationship between psychopathy and several relevant criteria was attenuated by faking.

Although there is no consensus regarding the impact of faking on validity or organizational decisions, organizations should not ignore the issue completely. The relationship of the variable of interest with other variables might not be noticeably impacted, but faking introduces systematic, construct-irrelevant variance into the measure, which undermines its construct validity (Tett & Simonet, 2011). It is hard to justify using an assessment for decision-making if it is unclear what it is measuring (Tett & Simonet, 2011).

Models of Faking

There has been a recent call to focus research on measuring faking using a theoretical model, rather than trying to piece together independent empirical studies (Griffith & Peterson, 2011). There are several models of faking that can be used as a basis for guiding faking research. An early model proposed by Snell, Sydell, and Lueke (1999) conceptualized an individual's tendency to fake as a combination of his/her motivation to fake and ability to fake. This relatively simple model provided a base for later models, such as McFarland and Ryan's (2000) model that described beliefs toward faking influencing intention to fake which influences faking behavior. They included ability to fake, situational influences, and opportunity to fake as moderators of this sequential relationship. Situational influences included factors such as desire for the job and presence of warnings explaining the consequences of being caught faking. Opportunity to fake refers to the ability of an individual to inflate his or her score. If the individual is already high on a positive trait (or low on a negative trait), then faking cannot have a large effect on his or her score. McFarland and Ryan (2006) expanded this model by including the components of the theory of planned behavior (attitudes, subjective norms, and perceived behavior control; Ajzen, 1991) as predictors of faking intentions.

In contrast to McFarland and Ryan's focus on beliefs and intentions, Tett and Simonet (2011) proposed a model that considers faking to be a type of performance. From their perspective, individuals who intend to fake are not all equally adept at increasing their score; some people may even make themselves look like a less desirable candidate by faking poorly. Conceptualizing faking as performance means that it can be defined as the product of a person's opportunity, ability, and motivation to fake, which implies that if there is a complete lack of opportunity, ability or motivation, then faking will not occur. This model offers utility for applied researchers because interventions created to target one of the three components can effectively reduce faking behavior.

Faking and dark personality. The majority of faking research has focused on the assessment of bright side personality traits, but faking is arguably an even larger concern for dark traits (Guenole, 2014). In general, individuals tend to de-emphasize their negative traits more than exaggerate positive characteristics (Donovan, Dwight, & Hurtz, 2003), which should lead them to deflate scores on dark traits more than they would inflate scores on bright side traits. In addition, based on Tett and Simonet's (2011) performance framework, people high in dark traits are more likely to fake based on their personality. They have the largest opportunity to fake because their true score is high, so they have more room to decrease their scores. Such individuals also have an increased motivation to fake because they have a natural propensity to self-enhance, and a personality test presents a trait-relevant situational cue that encourages expression of this self-enhancement propensity (Tett & Simonet, 2011). In addition, people high in Machiavellianism are especially likely to fake because a personality test in a selection situation presents a clear external reward (i.e., receiving a job offer), and the core of Machiavellianism is manipulation to receive a desired outcome. Finally, there is empirical

evidence that people high on Machiavellianism and psychopathy may have a higher ability to fake good (Book, Holden, Starzyk, Wasyliw, & Edwards, 2006; MacNeil, 2008).

Reducing Faking

There are three categories of methods that have been used in an attempt to control faking: demand reduction techniques, covariate techniques, and rational techniques (Paulhus & Vazire, 2007). Demand reduction techniques address the motivational aspect of faking by focusing on maximizing participant anonymity, which could include approaches such as the unmatched count technique and randomized response technique (see Coutts & Jann, 2011). This can be useful for basic research, but anonymity is clearly not possible in a selection or development context. Covariate techniques do not address motivation, ability, or opportunity directly. There are post hoc corrections that require administering a social desirability scale along with the measure of interest so that responses to the social desirability scale can be statistically controlled by partialling out that variance. However, this technique is not recommended because it removes construct-relevant variance that can actually reduce validity (Paulhus & Vazire, 2007). Paulus and Vazire (2007) recommend using rational techniques to prevent faking. Techniques such as warning instructions can address motivation by detailing the consequences of getting caught faking (e.g., removal from selection process) and cause respondents to question whether faking is beneficial. However, the effectiveness of warnings has not yet been adequately tested in applied settings, so it is unclear whether they would affect motivated job applicants in the same way as participants in a laboratory study (Stark, Chernyshenko, & Drasgow, 2011). In addition, Fisher and colleagues (2018) found that warning instructions have no effect for people high in psychopathy. In the same study, Fisher and colleagues (2018) found that a different rational technique, using a forced-choice response format, did decrease faking.

Forced-choice Method

The forced-choice method is another rational technique that involves presenting respondents with several stimuli simultaneously and requiring them to make comparative judgments about those stimuli (Brown, 2016). Multidimensional forced-choice (MFC) assessments are a specific type of the forced-choice method in which each item only measures one construct and blocks of items contain items representing different constructs rather than items measuring the same construct (Brown & Maydeu-Olivares, 2012). MFC assessments manage faking ability by preventing respondents from uniformly elevating responses across all dimensions, which makes it more difficult to fake good (Brown, 2016). For example, respondents could be presented items like: “Avoid danger,” “Have few close friends,” “Worry about things that might happen,” (from the forced-choice version of Goldsmiths 60-Item Questionnaire measuring maladaptive personality; Guenole et al. 2016) and be asked to report which is “most like me” and which is “least like me.” In this block, or collection of items presented together, respondents cannot give uniformly high or low responses. They must choose which item to endorse. Despite the potential for different response processes for comparative judgments, as opposed to the absolute judgment used on Likert-type responses, research has shown that response processes, scaling properties, and measurement precision are similar between single-stimulus and MFC assessments (Joubert, Inceoglu, Bartram Dowdeswell, & Lin, 2015; Sass, Frick, Reips, & Wetzel, 2018). In support of the MFC method, research has shown that score inflation due to faking is significantly reduced in MFC assessments as compared to a similar single-stimulus assessment (Christiansen, Burns, & Montgomery 2005).

The potential benefits of forced-choice measures have been known for decades, but MFC measures have not been widely adopted due to the psychometric problems with standard scoring

(Stark et al., 2005). When items are scored by summing the inverted rank order of items (e.g., in a 3-item block the trait ranked most like the individual gets 3 points for the item, the trait not chosen gets 2 points, and the trait rated least like the individual gets 1 point), then the test is said to be ipsative, meaning the set of responses always sum to the same total (Clemans, 1966). Ipsative data have several negative properties: interindividual comparisons are inappropriate, scale correlations are negative (regardless of true construct correlations), factor analysis is inappropriate, and criterion-related validity is distorted (Brown & Maydeu-Olivares, 2012; Clemans, 1966; Meade, 2004). Several item response theory (IRT) models have been developed that allow for non-ipsative scoring of MFC assessments; the most well-known of these models are Stark and colleagues' (2005) multi-unidimensional pairwise preference (MUPP) model and Brown and Maydeu-Olivares' (2011) Thurstonian IRT (TIRT).

MFC Scoring

The MUPP model estimates latent trait scores based on responses to paired items (Stark et al., 2005). This model alone is not enough to score a MFC assessment; it only estimates the latent trait scores for respondents, which means that an IRT model needs to be used to define item parameters in advance. This is usually done by gathering initial data on the items in a single-stimulus format and fitting an appropriate IRT model to estimate item parameters. Any IRT model appropriate for the data type can be used, and the mostly commonly chosen model used to determine the item parameters is the generalized graded unfolding model (GGUM; Roberts, Donoghue, & Laughlin, 2000).

The GGUM assesses an ideal point response process, which means that people endorse items that are close to their standing on the latent trait. Respondents can disagree with an item because it is located significantly above or below their true score on the latent trait. The

alternative to an ideal point process is a dominance response process, which assumes item endorsement represents a higher standing on the latent trait and is operationalized as a monotonic item response function. Dominance models are more parsimonious than ideal point models and better represent certain constructs. Research has found that a dominance model fits pathological personality better than an ideal point model (Carvalho, de Oliveira, Pessotto, & Vincenzi, 2015), thus dominance models are likely appropriate for measuring dark personality. Morillo and colleagues (2015) recently developed a dominance version of the MUPP (MUPP-2PL) model that uses the two-parameter logistic model instead of the GGUM to estimate item parameters. An advantage of this model is that it allows simultaneous estimation of item and person parameters, so the additional single-stimulus data required for the original MUPP is not needed.

The TIRT model is a dominance model that is described in terms of factor analysis. Although it can be equivalently described in IRT nomenclature, using factor analysis allows the model to be easily implemented by common structural equation modeling software, such as Mplus. Similar to the MUPP-2PL, this model can estimate both item and person parameters, so an additional data collection is not needed. The TIRT also has the advantage of easily accommodating blocks with more than two items. There is an extension of the original MUPP model that can estimate person parameters for larger blocks (Hontangas et al., 2015), but the MUPP-2PL can only be used with blocks of two items.

Choosing a model. Neither the MUPP, MUPP-2PL, nor TIRT model is inherently superior, so the choice of the model should be based on conceptual (dominance versus ideal point process) and pragmatic concerns (Brown, 2016). The MUPP model is the only option for ideal point processes, but additional resources are needed to estimate item parameters. If a dominance model is appropriate, then practical constraints can dictate whether to choose the

MUPP-2PL or TIRT. The MUPP-2PL is based in a Bayesian framework and uses a Markov Chain Monte Carlo (MCMC) estimation procedure, which can be computationally intensive (Morillo et al., 2015). However, the estimation procedure puts less constraints on the construction of the MFC assessment than either the MUPP or TIRT. In order to scale the latent metric, the authors of the MUPP model recommend blocks of unidimensional items with different location parameters (i.e., pair two items representing different levels of the same construct), and the authors of the TIRT model recommend opposite polarity (i.e., combine positively and reverse keyed items) blocks (Brown & Maydeu-Olivares, 2011; Stark et al., 2005). Both of these types of blocks will be difficult to match for social desirability, and discrepancies in their social desirability would make them easier to fake. Thus, if a MFC scale is not yet constructed, the MUPP-2PL may be the best scoring option. The advantages of the TIRT model are that it can be used for larger blocks, should not take as long to run as MCMC estimation used in the MUPP-2PL, and is easy to execute for individuals who are already familiar with structural equation modeling (Brown & Maydeu-Olivares, 2012). The TIRT model is likely the most accessible model for practical applications, especially if the MFC assessment is already written and construction of items is not a concern.

Current Paper

The purpose of the present paper is to examine whether a multidimensional forced-choice measure of dark personality is more resistant to faking than a single-stimulus assessment and determine its relationship with CWBs. The introduction of a faking resistant dark personality MFC assessment would be an important preliminary step toward utilizing knowledge regarding dark traits in organizations. Dark personality could be used to select people who are less likely to engage in CWBs, as well as inform the training and development of current employees.

In the only attempt to use an MFC assessment to measure maladaptive traits, Guenole and colleagues (2016) showed that an MFC scored with the TIRT model provided adequate fit to a measure of the trait model described in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) section on emerging models (APA, 2013), whereas the corresponding single-stimulus data had poor model fit. The DSM-5 trait model is an approach to psychiatric diagnosis that uses continuous scores on six personality dimensions, rather than conceptualizing people as fitting into a category (e.g. high on antagonism dimension rather than classified as having Narcissistic Personality Disorder; Guenole, 2014). This model has been theoretically related to the DT by conceptualizing the DSM-5 trait model as an overarching framework in which the DT traits are compound traits comprised of different levels of the six dimensions (Guenole, 2014). However, it is difficult to measure compound traits indirectly, and the DSM-5 trait model has very limited predictive validity evidence for organizational outcomes (Wille & De Fruyt, 2014). Therefore, I will create a new MFC assessment of the DT traits, which have an established relationship with CWBs (O'Boyle, et al., 2012). Based on previous findings of MFC measures of bright side personality traits (e.g. Christiansen et al., 2005), I predict that an MFC measure of the DT will be more resistant to faking than the corresponding single-stimulus measure.

Hypothesis 1: There will be less score deflation on an MFC measure of dark personality than a single-stimulus measure.

In addition, faking can attenuate the relationships between self-report measures and relevant criteria. Despite inconsistencies in the literature surrounding this issue (e.g., Peterson et al., 2011; Watts et al., 2016), in the case of dark personality measurement when respondents have the motivation to fake, the resulting score deflation should restrict the variance in the dark

personality measure, which would attenuate the relationship between dark personality and CWBs. Further, because an MFC measure should allow less faking, the relationship between an MFC assessment of dark traits and CWBs should be stronger than the relationship between a single-stimulus assessment of dark traits and CWBs.

Hypothesis 2: The relationship between dark personality and CWBs will be significantly weaker when respondents fake than when respondents answer honestly.

Hypothesis 3: In the presence of faking, an MFC measure of dark personality will have a stronger relationship with CWBs than a single-stimulus measure.

Finally, although the focus of this paper is dark personality, some authors have argued that the Honesty-Humility dimension of the HEXACO model is a viable way to measure the same constructs, just from the opposite pole (e.g., Lee & Ashton, 2014). Therefore, I will examine whether dark personality provides incremental validity over the HEXACO model for predicting CWBs.

Research Question: Does dark personality account for additional variance in CWBs beyond the HEXACO model?

Method

A pilot study was first conducted in order to develop both a single-stimulus and MFC measure of the DT. Details of this study are available in the appendix. Participants completed the DT assessments and the CWB measure at two different time points to reduce the influence of common method variance on the correlations between the DT traits and CWBs (Podsakoff & Organ, 1986).

Participants

One thousand and twelve workers from Amazon's Mechanical Turk completed the surveys at the first time point, and 764 returned for the second time point. Inclusion criteria required participants to be residents of the United States and work at least 30 hours per week. Participants received \$0.30 for completing the first time point and \$0.70 for completing the second time point. Several responses were removed for careless responding. Thirty-four participants who began the second time point missed an instructed response item (i.e. "Select once or twice for this item") and were automatically removed from the survey before they completed it. I also managed careless responding by calculating three post-hoc indices of careless responding: Mahalanobis Distance, Psychometric Synonyms, and Even Odd Consistency. Responses for each index were considered careless if the score on the index was more than two standard deviations from the mean. I only removed the data for participants who were flagged as an outlier by the Mahalanobis Distance and as inconsistent by either of the other indicators. This removed 25 responses. The analyses were conducted on the 705 remaining participants.

Sixty-one percent of the final sample was female. The participants had a mean age of 35.98 ($SD = 10.71$). Eighty-two percent of the sample identified as White, 8% as Black or African American, 5% as Asian and/or Asian American, and 5% identified as other races. Nine percent of the sample reported their ethnicity as Hispanic or Latino. Participants worked an average of 41.15 ($SD = 6.65$) hours per week.

Procedure

Data were collected at two time points (T1 & T2) with a minimum of a one-week time lag. At the first time point participants completed demographic items, as well as the MFC and

single-stimulus version of the dark personality assessment twice: once in an honest condition and once in a faking condition. Instructions for the faking condition asked participants to fake their responses as though they are applying for a job and want to look like an ideal job applicant (for similar instructions see Heggstad, Morrison, Reeve, & McCloy 2006; McFarland & Ryan, 2000). Participants completed the honest condition followed by the faking condition. Participants completed a short reaction to the assessment after each version of the assessment in the honest condition. The presentation of the MFC and single-stimulus assessment were randomized within condition. In order to create a strong incentive to comply with faking instructions, participants were informed that if they scored the highest in either faking condition they would receive a \$10 bonus.¹

Participants were sent an automated email with the link to the second part of the study, which contained CWB and HEXACO items, exactly one week after they completed T1. Participants who had not completed T2 were sent a reminder three days before collection ended and on the last day to complete the study. Then they completed the HEXACO items followed by a measure of self-reported CWBs for their current job.

Measures

In addition to the study measures, participants completed a six-item perceived overall justice scale (Ambrose & Schminke, 2009), a six-item job control scale (Jensen, Patel, & Messersmith, 2013), and a few items asking about their reaction to the dark personality assessments. The six reaction items were written for this study.

¹ They were also informed in the event of a tie the winner would be randomly selected from those with the highest score. Because there is no actual “ideal score,” every score tied, so everyone who completed both parts were given an equal chance to win the two \$10 bonuses. They were not informed of the opportunity to win a bonus until the faking instructions to ensure that the potential for a bonus would not increase motivation to fake in the honest condition.

Dark Personality. There were two versions of the dark personality scale. A multidimensional forced-choice version and a single-stimulus version using the same items as the MFC version. In a pilot study, I created both scales following Stark and colleagues (2005) outline of MFC scale development with a few modifications to include an assessment of content validity and dimensionality. This process included collecting a large pool of items, asking subject matter experts to rate the content validity of the items, asking a different panel of experts to rate the social desirability of the items, administering items to a large sample to establish the dimensionality of the DT traits, and pairing items based on social desirability. The details of the pilot study are described in the Appendix. For the MFC scale, participants were asked to choose which item was more true of their beliefs about themselves or others. The items can be found in Table 1. Participants responded to the single-stimulus version on a Likert-type scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

Bright Personality. Bright personality traits were measured with the HEXACO-60 (Ashton & Lee, 2009). The HEXACO-60 uses 10 items to measure each of the factors in the HEXACO model: Honesty-Humility, Emotionality, eXtraversion, Agreeableness, Conscientiousness, and Openness to experience. An example honesty-humility item is: “Having a lot of money is not especially important to me.” Items were rated on a scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). After running a confirmatory factor analysis (CFA) for the HEXACO items, I dropped one agreeableness item (“I am usually quite flexible in my opinions when people disagree with me”) and two conscientiousness items (“I always try to be accurate in my work, even at the expense of time” and “People often call me a perfectionist.”) because they had low factor loadings (i.e. < .35) onto their respective constructs. Cronbach’s

alpha for honesty-humility (.80), emotionality (.79), extraversion (.84), agreeableness (.77), conscientiousness (.80), and openness (.81) were all acceptable.

CWBs. Counterproductive work behaviors were assessed with the 32-item Counterproductive Work Behavior Checklist (CWB-C; Spector et al., 2006). Items were rated on a scale of 1 (*Never*) to 5 (*Every day*). I excluded three items from the analyses because they had low factor loadings onto the latent trait in a CFA. Each participant's CWB score was calculated as the mean of the 29 remaining items. The final scale had a high internal consistency reliability ($\alpha = .95$).

Results

A separate CFA was conducted for each DT trait using the honest single-stimulus responses to determine whether the three-dimensional factor structure (i.e. each DT trait as a separate, unidimensional construct) sample of 1,012 participants who completed the first time point. Four items had low factor loadings ($< .35$) on their respective constructs and were removed (see Table 1). All further analyses were conducted using this final version of the items. Model fit for the psychopathy model was good, but comparative fit indices for Machiavellianism and narcissism were relatively low (see Table 2). However, the Standardized Root Mean Square Residual was acceptable for both models ($\leq .08$; Hu & Bentler, 1999).

Dark Personality Scoring

The multidimensional forced choice measure was scored with both the MUPP-2PL and TIRT models. Both IRT models used the full sample of 1,012 participants for the estimation of item parameters. As described above, these models both have practical advantages in different contexts. I did not anticipate significantly different person parameters to be derived from either model because previous research indicates that both models result in similar person parameters

(Morillo et al., 2015), but both were included to demonstrate their utility in the context of scoring dark personality MFC assessments. The item parameters for the MUPP-2PL were estimated using an R routine obtained from Daniel Morillo (personal communication, March 19, 2018). This scoring uses MCMC estimation, which is an iterative resampling technique that uses random draws to obtain the sample distribution of interest.² The construct scores for faking and honest responses were then obtained using the honest item parameters so that they would both be on the same scale. Scores were computed using expected a posteriori estimation.

The TIRT model was scored in Mplus 8 using syntax generated by the Excel macro created by Brown and Maydeu-Olivares (2012). The honest scores were obtained from the original syntax. Similar to the MUPP-2PL scoring, the scores for the faking condition were obtained by estimating new factor scores while fixing the item parameters to the values found in the honest score estimation so that the scores were on the same scale.

An IRT model was used for the single-stimulus data to parallel the scoring of the MFC measure. Similar conclusions would likely be drawn using typical classical test theory scoring procedures because latent trait scores from IRT models are very highly correlated with the more frequently used classical test theory scoring (Embretson & Reise, 2000). The single-stimulus version of the dark personality measure was scored with a dominance IRT model appropriate for Likert-type responses: the graded response model (Samejima, 1969). The model parameters were also created using the full sample of 1,012 participants. The appropriateness of the model was assessed using the M_2 statistic (Maydeu-Olivares & Joe, 2006) and the $S-\chi^2$ (Orlando & Thissen,

² Each series of random draws is referred to as a chain. I specified four chains each with 500,000 draws. The first 200,000 were burn-in, and only every 50th draw was saved. This yielded 4 (chains) X 300,000 / 50 (saved) = 24,000 total saved draws to create the posterior distribution. Gelman and Rubin's (1992) statistic for all parameters were below the cut off of 1.2, which indicates that the chains had converged. (For more information on MCMC and other Bayesian estimation see Kruschke, Aguinis, and Joo, 2012.) These item parameters were obtained using honest responses.

2000) statistic for each item (Tables 3-5). Non-significant p -values indicate good fit for both of these indices. Although all models had significant M_2 values, significance tests have more power to detect differences with a large sample size. Fit at the item level using the $S-\chi^2$ showed stronger support for the appropriateness of the model. Very few items were found to fit poorly for the honest responses.³

Descriptive statistics and zero-order correlations of all dark personality scores, the six HEXACO personality factors, and CWBs can be found in Tables 6-8. In order to determine whether participants faked their responses on each measure, I conducted t -tests for dependent means to look for significant increase in scores between honest and faking conditions (see Table 9). The t -tests were significant for all three scoring methods (MUPP-2PL MFC, TIRT MFC, and graded response model single-stimulus). This indicates that there was faking via score inflation and deflation. Machiavellianism and psychopathy both showed significant deflations, whereas narcissism showed significant inflation, as is consistent with a “fake-good” approach.

Test of Hypotheses

The first hypothesis stated that there would be less score deflation on the MFC measure than on single-stimulus measure. This hypothesis was tested by conducting t -tests for dependent means to examine differences in faking scores for the single-stimulus versus MFC formats.

Before the t -tests, scores for all three models (MUPP-2PL MFC scores, TIRT MFC scores, and graded response model single-stimulus) were standardized to be expressed in standard deviation

units of their respective honest scores (i.e., $\frac{x_{faking} - \bar{x}_{honest}}{SD_{honest}}$). The standardized score then

represents the amount of faking with a larger magnitude indicating more faking. The sign

³ Many items fit poorly for the faking responses. This is expected because the faking responses were scored with item parameters from the honest responses. If the applicant instructions successfully changed the way participants responded, then the faking responses should not fit the model for honest responses well.

indicates the direction of faking, so negative standardized scores represent more faking because respondents lowered their scores on undesirable traits. These standardized scores for the MUPP-2PL MFC scoring and TIRT MFC scoring were compared to the single-stimulus graded response model scores using *t*-tests for dependent means. The standardized MFC psychopathy and Machiavellianism (except MUPP-2PL) scores were significantly lower than the standardized graded response model scores, and the standardized MFC narcissism scores were significantly higher than the standardized single-stimulus scores (Table 10). The differences for psychopathy and Machiavellianism were in the opposite direction I hypothesized, and although the difference in narcissism was in the hypothesized direction, examination of the single-stimulus scores indicate that participants sought to increase their narcissism scores. Thus, the MFC scores showed more faking than the single-stimulus scores. This provides evidence against Hypothesis 1.

The second hypothesis asserted that the relationship between CWBs and dark personality would be stronger in the honest condition than the faking condition. I tested this hypothesis by computing correlations between each DT trait and CWBs and comparing them using the method for comparing dependent, overlapping correlations described by Wilcox (2009). I used this method because it accounts for nonnormality and heteroskedasticity, which are present in the CWB data. I used the R script provided by Rand R. Wilcox on his website (<https://dornsife.usc.edu/labs/rwilcox/software/>) to compute confidence intervals for the differences between the correlations (see Table 11). The confidence intervals for narcissism for the MFC scoring methods and all three DT traits for the single-stimulus scoring method did not contain zero, which provided partial support for the hypothesis.

The third hypothesis stated that the MFC scores would have a stronger relationship with CWBs than the single-stimulus scores when respondents are faking. I tested this hypothesis in the same way as Hypothesis 2 (see Table 12). The confidence intervals for the differences all contained zero for Machiavellianism and psychopathy, which does not provide support for the hypothesis. The confidence interval for narcissism did not contain zero when comparing the single-stimulus to the TIRT scoring, which indicated a significant difference between the correlations between MFC scoring and CWBs and single-stimulus scoring and CWBs. The magnitude of the MFC correlations was larger, but the correlations were negative instead of positive. This indicates that higher MFC narcissism scores were associated with fewer CWBs. Thus, Hypothesis 3 is not supported.

Research Question

The research question asked whether dark traits explained incremental variance in CWBs beyond the HEXACO model. I addressed this question using hierarchical regressions for the honest single-stimulus DT score and the two honest MFC DT scores. Only the honest condition responses were used to provide a more direct comparison. A significant difference between models was assessed using a Wald test with a heteroskedasticity consistent covariance matrix. I used the HC3 heteroskedasticity consistent covariance matrix (MacKinnon & White, 1985), as recommended by Long and Ervin (2000), to account for the heteroskedasticity in the CWB scores. For each set of regressions all six HEXACO scale scores were entered in the first step, and the DT scores were entered in the second step (see Table 13). The research question was answered affirmatively for both MFC and single-stimulus scores.

Discussion

The present study administered a newly developed MFC and parallel single-stimulus measure of the DT and assessed their resistance to faking. There was significant faking on the MFC and single-stimulus versions of the assessment which indicates that participants were able to significantly decrease their scores on psychopathy and Machiavellianism. Counter to my expectations, participants seemed to purposefully increase their narcissism scores when asked to respond like an ideal job applicant. This shows the importance of considering recent calls to not overlook the benefits of dark traits (e.g., Judge & LePine, 2007; Smith, Hill, Wallace, Recendes, & Judge, 2018). Although all dark traits potentially have contexts in which they could be beneficial, participants only increased scores for narcissism. This preference could indicate a perception that narcissism has more generally applicable benefits than the other two DT traits. It could also reflect the more specific findings that people higher in narcissism are rated positively in interviews (Paulhus, Westlake, Calvez, & Harms, 2013). Participants could have recognized this trend and thus increased their narcissism scores in order to match what they believed employers want, rather than reflecting a personal belief that narcissism is a desirable trait. Future research should determine whether this preference for narcissism reflects a belief that narcissism is related to effective performance or a belief that employers tend to hire people who are more narcissistic.

The first hypothesis that there would be less score deflation on the MFC measure was not supported. The difference between the single-stimulus and MFC scores for Machiavellianism and psychopathy were significantly inflated instead of deflated, indicating that there was significantly *more faking* on the MFC measure than the single-stimulus measure. In addition, although the narcissism scores were significantly less deflated on the MFC measure, the increase

in scores from honest to the faking condition in the single-stimulus assessment indicates that participants were purposefully trying to increase rather than decrease their narcissism scores. Under the assumption that respondents wanted to increase narcissism scores, the results showed that the MFC assessment allowed for significantly more faking. Thus, rather than decreasing faking, the MFC format allowed for increase faking on all three DT traits. There are several possible explanations for this finding.

First, the forced-choice method works best when one item is not clearly “better,” and if narcissism is a more valued construct than psychopathy and Machiavellianism, then people may have uniformly chosen narcissism items when given the chance, which would undermine the advantage of the forced-choice method. If participants always chose narcissism items, then pairs involving narcissism items would consistently lead to the choice of narcissism items as the preferred item. This in turn would deflate Machiavellianism and psychopathy scores. However, as can be seen in Table 14, respondents did not always choose a narcissism item when given a chance. Many items were skewed such that narcissism was chosen by more participants, but a substantial number of participants also chose the psychopathy or Machiavellianism option.

Another potential explanation is that there may be a difference between what is perceived as socially desirable in general and what is perceived to be desirable for an employee. For example, the item “I am willing to be unethical if I believe it will help me succeed” may be seen as desirable for employees, but undesirable in a broader societal context.

Further, which traits are desirable may vary by specific job. Research has shown that people who work in different work sectors have different patterns of dark side traits (Furnham, Hyde, & Trickey, 2014), so beliefs about what level of a dark trait is ideal likely also vary by job. For example, in the pilot study (see Appendix) endorsing the item “People who are stupid

enough to get ripped off usually deserve it” had an average subject matter expert social desirability rating of 1.17 on scale of 1 to 5, so it is generally low in social desirability. However, if a participant was thinking of applying for a sales position, then endorsing this item may be seen as highly desirable to be an effective salesperson. In other occupations, this may not be seen as job relevant, so ratings would be closer to the general social desirability ratings (e.g. “ripping people off” is not relevant for the job tasks of a teacher, so endorsing the item would just point to the applicant being an unethical person). A limitation of this research is not asking participants what job they had in mind when responding. Future research should compare the effectiveness of MFC scales developed with general social desirability ratings, employment-specific desirability ratings, and job-specific social desirability ratings.

The second hypothesis received mixed support. Honest responses were significantly more related to CWBs for all traits in the single-stimulus measure, and honest responses for narcissism were significantly more related to CWBs for the MFC measure. The results for the single-stimulus measure were as anticipated. The magnitude of the relationship between the DT traits and CWBs decreased significantly for all three traits, and for narcissism the decrease was so substantial that there was no longer a significant relationship. However, I anticipated that this difference would be due to a restriction of variance in the scores from the faking condition, but the results in Table 15 indicate this was not the case. Levene’s test (1960) using the median as the measure of central tendency indicated that there was actually significantly more variance in the faking single-stimulus scores than the honest scores. This indicates that participants had different ideas about what kinds of responses would be appropriate for an ideal job applicant. Different ideas about the definition of an ideal applicant could be due to beliefs about a good employee versus a good person or the characteristics of a good employee for a specific type of

position, as discussed above. Differences in these beliefs could also be due to more stable individual differences. This would be consistent with Tett and Simonet's (2011) model that considers faking a type of performance. Some individual difference variables could be related to a higher ability to fake effectively.

In order to assess this potential explanation, I conducted an exploratory analysis of individual differences influencing scores in the faking condition by regressing faking scores on HEXACO scores and including the relevant honest score as a control (Table 16). Robust standard errors (HC3 method; MacKinnon & White, 1985) were used to account for heteroskedasticity. Results showed that extraversion is significantly inversely related to narcissism faking scores. There is conceptual overlap between narcissism and extraversion, so a relationship between extraversion and belief about ideal level of narcissism is not surprising. Indeed, in this sample honest narcissism scores for MUPP-2PL ($r = .38$), TIRT ($r = .38$), and single-stimulus ($r = .39$) scoring were all significantly (p 's $< .01$) positively related to extraversion. The inverse relationship between extraversion and faking narcissism scores could be a reaction to feedback that people high and low in extraversion have received. For example, someone low in extraversion may have been told that he or she needs to be more outgoing, so he or she would find items such as "I am apt to show off if I get the chance" more desirable, whereas people who are high in extraversion may have been reacting to the opposite type of feedback.

There were some other relationships between personality and faking scores. Honesty-Humility predicted Machiavellianism and psychopathy scores in the faking condition, but only for the single-stimulus scoring. In addition, openness was positively related to faking narcissism scores and significantly negatively related to psychopathy faking scores (Table 16). There is not

a strong theoretical basis for these relationships, so future research should determine whether they can be reproduced and if so, why they should occur.

The results for the second hypothesis using the MFC narcissism scores deserve special attention. The relationship between MFC narcissism scores changed from significant and positive for honest responses to significant and negative for the faking responses in all but the MUPP-2PL final items score. Narcissism has an established positive relationship with CWBs (e.g. O'Boyle et al., 2012), so the reversal of this relationship could indicate that the construct the assessment measures changes when people are responding as an ideal job applicant. Perhaps the items are viewed more as a more positive trait, so endorsing the "positive narcissism" predicts less CWBs. This reconceptualization of narcissism items as positive would explain the finding that participants purposefully increased scores on the narcissism items in the faking condition. Regardless of the cause, this reverse relationship has important implications for practitioners who measure narcissism in contexts where respondents are likely to fake, such as selection. If respondents fake, then narcissism scores may have a negative relationship with CWBs, so using these kinds of assessments in hiring decisions would actually increase the probably that hired applicant would engage in of CWBs.

The third hypothesis, that the MFC scores would have a stronger relationship with CWBs than the single-stimulus scores when respondents are faking, was not supported. The differences in the relationships were nonsignificant for Machiavellianism and psychopathy. The differences for narcissism were significant (except for MUPP-2PL scoring), but the correlations were negative. Based on previous research (e.g., O'Boyle et al., 2012), I anticipated a positive correlation between narcissism and CWBs, so finding a significantly stronger negative correlation provides evidence against the hypothesis. This further demonstrates the importance of

assessing an MFC assessment's resistance to faking before implementing it. Using MFC scores that have the opposite relationship with criteria would be more detrimental in applied settings than the lack of relationship found in the single-stimulus scores.

The lack of support for hypothesis 3 is not surprising in light of the finding that MFC scores did not have less deflation/inflation than single-stimulus scores (lack of support for hypothesis 1). I anticipated a stronger correlation because I expected the faking scores on the MFC measure to be closer to the honest scores (i.e. more accurate) than the single-stimulus scores, and a more accurate score would likely be more related to a theoretically related construct. Future research should test this hypothesis with an MFC measure that effectively reduces faking.

The research question was answered affirmatively for the final items; DT traits account for incremental variance in CWBs beyond HEXACO model. Further examination of where explanatory power was added showed that of the DT traits, only psychopathy slopes were statistically significant when controlling for all of the other predictors (and no DT slopes were significant using TIRT scoring). Thus, psychopathy may be a more important predictor of CWBs than narcissism and Machiavellianism when HEXACO factors are already accounted for. It should be noted that the HEXACO measure was given at the second time point, which was the same time as the CWB items, but DT items were given at the first time point. Administering the HEXACO and CWB measures at the same time increased the likelihood of common method variance (Podsakoff & Organ, 1986), which could have inflated the correlation between the two measures. Thus, this was a conservative assessment of the incremental importance of DT traits.

Limitations and Future Directions

A limitation of this study was the focus on the DT as a set of dark traits. I chose to measure the DT because there is not an alternative comprehensive taxonomy of dark traits, and the DT has been shown to predict relevant organizational outcomes (i.e. CWBs). Although these are the most well-known and studied dark traits, they were not developed in the same holistic way as bright-side taxonomies (e.g. FFM & HEXACO). Each of the DT traits was studied individually first, and Paulhus and Williams (2002) studied them in combination in order to disentangle three conceptually similar constructs. The DT traits were combined and labeled because they are more effectively studied together in order to delineate the degree of distinctiveness and overlap (Paulhus, 2014). Grouping similar traits does not guarantee the comprehensive coverage of a dark domain. In fact, recent research has added “everyday sadism” to the DT and advocated studying this new “Dark Tetrad” (Buckels, Jones, & Paulhus, 2013). In addition, these groups of dark traits explicitly only include traits that share a core of callousness (Paulhus, 2014), which precludes the inclusion of other potentially important constructs.

Other frameworks based on the DSM-IV personality disorders (i.e., the Hogan Developmental Survey; Hogan & Hogan, 2009) and DSM-5 trait model (Guenole, 2014) have some support. However, reliance on subclinical disorders also does not guarantee comprehensive coverage. Additionally, although measuring subclinical levels of disorders has not been problematic to date, there are potential concerns that a measure interpreted as a clinical assessment, rather than an assessment designed for normal populations, would be considered a preemployment medical examination that violates the Americans with Disability Act (Wu & LeBreton, 2011). The assessment of maladaptive personality-based DSM-5 trait model (Guenole et al. 2016) may be the most comprehensive view of maladaptive traits; however, it classifies all

of the DT traits as different levels of one factor (antagonism). This clustering is precisely what Paulhus and Williams (2002) sought to avoid by studying these traits simultaneously. In addition, research on this model is relatively new, so there is no established relationship between the trait model and workplace behaviors. Future research should examine these potential relationships and the viability of this model as a framework for dark personality.

Conclusions

This study demonstrated the importance of assessing the resistance to faking for MFC assessments before implementation. I created an MFC measure of the DT using steps suggested by Stark and colleagues (2005) but found that the measure was not more resistant to faking than a single-stimulus version. The implementation of this assessment in an applied context, such as selection, would be at best no more informative than a simpler single-stimulus assessment and at worst it would provide misinformation. Use of MFC narcissism scores, if applicants were faking, could actually increase the likelihood of hiring candidates who are *more* likely to engage in CWBs. Future research should investigate the boundary conditions, such as type of construct measured and testing situation, for when MFC assessments are most likely to be resistant to faking.

REFERENCES

- Ackerman, R. A., Witt, E. A., Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., & Kashy, D. A. (2011). What does the narcissistic personality inventory really measure? *Assessment, 18*(1), 67-87. 10.1177/1073191110382845
- Ajzen, I. (1991). Theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179-211.
- Ambrose, M. L., & Schminke, M. (2009). The role of overall justice judgments in organizational justice research: A test of mediation. *Journal of Applied Psychology, 94*, 491-500. doi:10.1037/a0013203
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Amernic, J. H., & Craig, R. J. (2010). Accounting as a facilitator of extreme narcissism. *Journal of Business Ethics, 96*, 79–93.
- 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. doi:10.1037/0021-9010.92.2.410
- Boddy, C. R. P. (2010). Corporate psychopaths and organizational type. *Journal of Public Affairs, 10*, 300–312.
- Book, A. S., Holden, R. R., Starzyk, K. B., Wasylkiw, L., & Edwards, M. J. (2006). Psychopathic traits and experimentally induced deception in self-report assessment.

Personality and Individual Differences, 41(4), 601–608.

<http://dx.doi.org/10.1016/j.paid.2006.02.011>

- Brown, A. (2016). Item response models for forced-choice questionnaires: A common framework. *Psychometrika*, 81(1), 135-160. doi:10.1007/s11336-014-9434-9
- Brown, A., & Maydeu-Olivares, A. (2011). Item response modeling of forced-choice questionnaires. *Educational and Psychological Measurement*, 71, 460–502.
- Brown, A., & Maydeu-Olivares, A. (2012). Fitting a Thurstonian IRT model to forced-choice data using Mplus. *Behavior Research Methods*, 44, 1135–1147.
- Buckels, E. E., Jones, D. N., & Paulhus, D. L. (2013). Behavioral confirmation of everyday sadism. *Psychological Science*, 24, 2201-2209. doi:10.1177/0956797613490749
- Burke, M. J., & Dunlap, W. P. (2002). Estimating interrater agreement with the average deviation index: A user's guide. *Organizational Research Methods*, 5(2), 159-172. doi:10.1177/1094428102005002002
- Carvalho, L. d. F., de Oliveira, A. Q., Pessotto, F., & Vincenzi, S. L. (2015). Application of the unfolding model to the aggression dimension of the dimensional clinical personality inventory (IDCP). *Revista Colombiana de Psicología*, 23, 339-349. doi:10.15446/rcp.v23n2.41428
- Christiansen, N. D., Burns, G. N., & Montgomery, G. E. (2005). Reconsidering forced-choice item formats for applicant personality assessment. *Human Performance*, 18, 267-307. doi:10.1207/s15327043hup1803_4
- Clemans, W. V. (1966). *An analytical and empirical examination of some properties of ipsative measures* (Psychometric Monograph No. 14). Retrieved from <https://www.psychometricsociety.org/sites/default/files/pdf/MN14.pdf>

- Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin, 136*, 1092-1122. doi:10.1037/a0021212
- Cooke D. J., Michie C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment, 13*, 171–188.
- Costa, J., P T, & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the revised NEO personality inventory. *Journal of Personality Assessment, 64*(1), 21-50. doi:10.1207/s15327752jpa6401_2
- Costa, P. T., McCrae, R. R., & Dye, D. A. (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO personality inventory. *Personality and Individual Differences, 12*(9), 887-898. doi:10.1016/0191-8869(91)90177-D
- Coutts, E., & Jann, B. (2011). Sensitive Questions in Online Surveys: Experimental Results for the Randomized Response Technique (RRT) and the Unmatched Count Technique (UCT). *Sociological Methods & Research, 40*. doi:10.1177/0049124110390768.
- Dahling, J. J., Whitaker, B. G., & Levy, P. E. (2009). The development and validation of a new Machiavellianism scale. *Journal of Management, 35*, 219-257.
doi:10.1177/0149206308318618
- Donovan, J. J., Dwight, S. A., & Hurtz, G. M. (2003). An assessment of the prevalence, severity, and verifiability of entry-level applicant faking using the randomized response technique. *Human Performance, 16*, 81-106. doi:10.1207/S15327043HUP1601_4
- Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology, 37*, 90-93.
doi:10.1037/h0058073

- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, N.J: L. Erlbaum Associates.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299.
- Fisher, P. A., Robie, C. Christiansen, N. S., & Komar, S. (2018): The impact of psychopathy and warnings on faking behavior: A multisaturation perspective. *Personality and Individual Differences, 127*, 39-34. <https://doi.org/10.1016/j.paid.2018.01.033>
- Furnham, A., Hyde, G., & Trickey, G. (2014). Do your dark side traits fit? dysfunctional personalities in different work sectors: Dark side traits. *Applied Psychology, 63*, 589-606. doi:10.1111/apps.12002
- Gelman, A. E., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science, 7*, 457-472. doi:10.1214/ss/117701113
- Greenberg, J. (2002). Who stole the money, and when? individual and situational determinants of employee theft. *Organizational Behavior and Human Decision Processes, 89*(1), 985-1003. doi:10.1016/S0749-5978(02)00039-0
- Griffith, R. L., & Peterson, M. H. (2011). One piece at a time: The puzzle of applicant faking and a call for theory. *Human Performance, 24*(4), 291-301. doi:10.1080/08959285.2011.597474
- Guenole, N. (2014). Maladaptive personality at work: Exploring the darkness. *Industrial and Organizational Psychology, 7*(1), 85-97. 10.1111/iops.12114

- Guenole, N., Brown, A., A., Cooper, A., J. (2016): Forced-choice assessment of work-related maladaptive personality traits: Preliminary evidence from an application of Thurstonian item response modeling. *Assessment*, 1-13. doi:10.1177/1073191116641181
- Harms, P. D., & Spain, S. M. (2015). Beyond the bright side: Dark personality at work. *Applied Psychology*, 64(1), 15-24. 10.1111/apps.12042
- Heggestad, E. D., Morrison, M., Reeve, C. L., & McCloy, R. A. (2006). Forced-choice assessments of personality for selection: Evaluating issues of normative assessment and faking resistance. *Journal of Applied Psychology*, 91(1), 9-24. doi:10.1037/0021-9010.91.1.9
- Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology*, 92, 1270-1285. doi:10.1037/0021-9010.92.5.1270
- Hogan, R., & Hogan, J. (2009). Hogan Development Survey manual (3rd Ed.). Tulsa, OK: Hogan Assessment Systems.
- Hontangas, P. M., de la Torre, J., Ponsoda, V., Leenen, I., Morillo, D., & Abad, F. J. (2015). Comparing traditional and IRT scoring of forced-choice tests. *Applied Psychological Measurement*, 39, 598-612. doi:10.1177/0146621615585851
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 32, 179-185.
- Hu L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives, *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1-55, doi:10.1080/10705519909540118

- Jensen, J. M., Patel, P. C., & Messersmith, J. G. (2013). High-performance work systems and job control: Consequences for anxiety, role overload, and turnover intentions. *Journal of Management*, *39*, 1699-1724. doi:10.1177/0149206311419663
- Jonason, P. K., & Webster, G. D. (2010). The Dirty Dozen: A concise measure of the dark triad. *Psychological Assessment*, *22*, 420–432
- Jonason, P. K., Wee, S., & Li, N. P. (2014). Thinking bigger and better about “Bad apples”: Evolutionary Industrial–Organizational psychology and the dark triad. *Industrial and Organizational Psychology*, *7*(1), 117-121. 10.1111/iops.12118
- Jones, D. N., & Paulhus, D. L. (2014). Introducing the short dark triad (SD3): A brief measure of dark personality traits. *Assessment*, *21*(1), 28-41. doi:10.1177/1073191113514105
- Joubert, T., Inceoglu, I., Bartram, D., Dowdeswell, K., & Lin, Y. (2015). A comparison of the psychometric properties of the forced choice and likert scale versions of a personality instrument: Psychometric properties of forced choice and likert rating scales. *International Journal of Selection and Assessment*, *23*, 92-97. doi:10.1111/ijsa.12098
- Judge, T. A., & LePine, J. A. 2007. The bright and dark sides of personality: Implications for personnel selection in individual and team contexts. In J. Langan-Fox, C. L. Cooper, & R. J. Klimoski (Eds.), *Research companion to the dysfunctional workplace*: 332-355. Northampton, MA: Edward Elgar.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151. doi:10.1177/001316446002000116

- 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*, 140-154. doi:10.1037/0021-9010.93.1.14
- König, C. J., Steiner Thommen, L. A., Wittwer, A., & Kleinmann, M. (2017). Are observer ratings of applicants' personality also faked? Yes, but less than self-reports. *International Journal of Selection and Assessment, 25*, 183-192. doi:10.1111/ijsa.12171
- Kruschke, J. K., Aguinis, H., & Joo, H. (2012). The time has come: Bayesian methods for data analysis in the organizational sciences. *Organizational Research Methods, 15*(4), 722-752. doi:10.1177/1094428112457829
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*, 563-575. doi:10.1111/j.1744-6570.1975.tb01393.x
- Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO Personality Inventory. *Multivariate Behavioral Research, 39*, 329–358.
- 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. doi:10.1207/s15327043hup1802_4
- Levene, H. (1960). Robust tests for equality of variances. In I. Olkin (Ed.), *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling* (pp. 278–292). Stanford University Press.
- Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology, 68*, 151-158. Doi:10.1037/0022-3514.68.1.151

- Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, *54*, 217-224.
doi:10.1080/00031305.2000.10474549
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.
- Machiavelli, N. (1950). *The Prince*. New York, NY: Modern Library. (Original work published 1532)
- MacNeil, B. M. (2008). *The Dark Triad and faking ability on self-report personality inventories and autobiographical accounts* (Unpublished doctoral dissertation) Kingston, Ontario: Queen's University. <http://hdl.handle.net/1974/1558>
- Mahmut, M. K., Menictas, C., Stevenson, R. J., & Homewood, J. (2011). Validating the factor structure of the self-report psychopathy scale in a community sample. *Psychological Assessment*, *23*, doi:670-678. 10.1037/a0023090
- Martinko, M. J., Gundlach, M. J., Douglas, S. C., (2002). Toward and integrative theory of counterproductive workplace behavior: A causal reasoning perspective. *International Journal of Selection and Assessment*, *10*, 36-50.
- Maydeu-Olivares, A., & Joe, H. (2006). Limited information goodness-of-fit testing in multidimensional contingency tables. *Psychometrika*, *71*, 713-732.
- McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*(2), 175-215. doi:10.1111/j.1467-6494.1992.tb00970.x
- McFarland, L. A., & Ryan, A. M. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology*, *85*, 812-821. doi:10.1037/0021-9010.85.5.812

- McFarland, L. A., & Ryan, A. M. (2006). Toward an integrated model of applicant faking behavior. *Journal of Applied Social Psychology, 36*(4), 979-1016. doi:10.1111/j.0021-9029.2006.00052.x
- Meade, A. W. (2004). Psychometric problems and issues involved with creating and using ipsative measures for selection. *Journal of Occupational and Organizational Psychology, 77*, 531-552. doi:10.1348/0963179042596504
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Morillo, D., Leenen, I., Abad, F. J., Hontangas, P., de la Torre, J., & Ponsoda, V. (2016). A dominance variant under the multi-unidimensional pairwise-preference framework: Model formulation and Markov Chain Monte Carlo estimation. *Applied Psychological Measurement, 40*, 500-516. doi:10.1177/0146621616662226
- Mueller-Hanson, R., Heggestad, E. D., & Thornton, G. C., III. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. *Journal of Applied Psychology, 88*, 348-355. doi:10.1037/0021-9010.88.2.348
- Muris, P., Merckelbach, H., Otgaar, H., & Meijer, E. (2017). The malevolent side of human nature: A meta-analysis and critical review of the literature on the dark triad (narcissism, Machiavellianism, and psychopathy). *Perspectives on Psychological Science, 12*(2), 183-204. doi:10.1177/1745691616666070
- O'Boyle, E. H., Forsyth, D. R., Banks, G. C., & McDaniel, M. A. (2012). A meta-analysis of the Dark Triad and work behavior: A social exchange perspective. *Journal of Applied Psychology, 97*, 557-579.

- Oh, I., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *The Journal of Applied Psychology, 96*(4), 762-773. doi:10.1037/a0021832
- 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. doi:10.1037/0021-9010.81.6.660
- Orlando, M., & Thissen, D. (2000). Likelihood-based item fit indices for dichotomous item response theory models. *Applied Psychological Measurement, 24*, 50-64.
- Paulhus, D. L. (2014). Toward a taxonomy of dark personalities. *Current Directions in Psychological Science, 23*(6), 421-426. doi:10.1177/0963721414547737
- Paulhus, D. L., Neumann, C. S., & Hare, R. D. (2009). *Manual for the Hare Self-Report Psychopathy Scale*. Toronto, Ontario, Canada: Multi-Health Systems
- Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 224-239). New York, NY: Guilford Press
- Paulhus, D. L., Westlake, B. G., Calvez, S. S., & Harms, P. D. 2013. Self-presentation style in job interviews: The role of personality and culture. *Journal of Applied Social Psychology, 43*, 2042-2059.
- 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*, 270-290. doi:10.1080/08959285.2011.580808
- Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management, 12*(4), 531-544. doi:10.1177/014920638601200408

- Raskin, R. N., & Hall, C. S. (1979). A Narcissistic Personality Inventory. *Psychological Reports*, 45, 590.
- Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics*, 4, 207–230.
doi:10.3102/10769986004003207
- Rogoza, R., & Cieciuch, J. (2018). Dark triad traits and their structure: An empirical approach. *Current Psychology*, Advanced online publication. doi:10.1007/s12144-018-9834-6
- Roberts, J. S., Donoghue, J. R., & Laughlin, J. E. (2000). A general item response theory model for unfolding unidimensional polytomous responses. *Applied Psychological Measurement*, 24, 3–32.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48, 1-36.
- Sackett, P. R. (2011). Integrating and prioritizing theoretical perspectives on applicant faking of personality measures. *Human Performance*, 24(4), 379-385.
doi:10.1080/08959285.2011.597478
- Sackett, P. R., Berry, C. M., Wiemann, S. A., & Laczko, R. M. (2006). Citizenship and counterproductive behavior: Clarifying relations between the two domains. *Human Performance*, 19(4), 441-464. 10.1207/s15327043hup1904_7
- Sackett, P. R., & DeVore, C. J. (2001). Counterproductive behaviors at work. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work, and organizational psychology* (Vol. 1, pp. 145–164). London: Sage.
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, 34, 100–114.

- Sass, R., Frick, S., Reips, U., & Wetzel, E. (2018): Taking the test taker's perspective: Response process and test motivation in multidimensional forced-choice versus rating scale instruments. *Assessment*, 1-13. <https://doi.org/10.1177/1073191118762049>
- Schyns, B. (2015). Dark personality in the workplace: Introduction to the special issue. *Applied Psychology*, 64(1), 1-14. 10.1111/apps.12041
- Smith, M. B., Hill, A. D., Wallace, J. C., Recendes, T., & Judge, T. A. (2018). Upsides to dark and downsides to bright personality: A multidomain review and future research agenda. *Journal of Management*, 44(1), 191-217. doi:10.1177/0149206317733511
- Snell, A. F., Sydell, E. J., & Lueke, S. B. (1999). Towards a theory of applicant faking: Integrating studies of deception. *Human Resource Management Review*, 9, 219-242. doi:10.1016/S1053-4822(99)00019-4
- Spain, S. M., Harms, P., & LeBreton, J. M. (2014). The dark side of personality at work. *Journal of Organizational Behavior*, 35, S41-S60. 10.1002/job.1894
- Spector P. E., & Fox S. (2005). The stressor-emotion model of counterproductive work behavior. In Fox S, Spector PE (Eds.), *Counterproductive work behavior: Investigations of actors and targets* (pp. 151–174). Washington, DC: American Psychological Association.
- Spector, P. E., Fox, S., Penney, L. M., Bruursema, K., Goh, A., & Kessler, S. (2006). The dimensionality of counterproductivity: Are all counterproductive behaviors created equal? *Journal of Vocational Behavior*, 68, 446–460. doi:10.1016/j.jvb.2005.10.005
- Stark, S., Chernyshenko, O. S., & Drasgow, F. (2005). An IRT approach to constructing and scoring pairwise preference items involving stimuli on different dimensions: The multi-dimensional pairwise-preference model. *Applied Psychological Measurement*, 29(3), 184-203. doi:10.1177/0146621604273988

- Stark, S., Chernyshenko, O. S., & Drasgow, F. (2011). Constructing fake-resistant personality tests using item response theory: High stakes personality testing with multidimensional pairwise preferences. In M. Ziegler, C. MacCann, & D.R. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 214-239). New York, NY: Oxford University Press.
- Tett, R. P., & Simonet, D. V. (2011). Faking in personality assessment: A "multisaturation" perspective on faking as performance. *Human Performance*, *24*(4), 302.
doi:10.1080/08959285.2011.597472
- Usami, S. Sakamoto, A, Naito, J., & Abe, Y. (2016). Developing pairwise preference-based personality test and experimental investigation of its resistance to faking effect by item response model, *International Journal of Testing*, *16*, 288-309,
doi:10.1080/15305058.2016.1145123
- Watts, A. L., Lilienfeld, S. O., Edens, J. F., Douglas, K. S., Skeem, J. L., Verschuere, B., & LoPilato, A. C. (2016). Does response distortion statistically affect the relations between self-report psychopathy measures and external criteria? *Psychological Assessment*, *2*, 294-306. doi:10.1037/pas0000168
- Wilcox, R. R. (2009). Comparing pearson correlations: Dealing with heteroscedasticity and nonnormality. *Communications in Statistics- Simulation and Computation*, *38*(10), 2220-2234. doi:10.1080/03610910903289151
- Wille, B., & Fruyt, F. D. (2014). Fifty shades of personality: Integrating Five-Factor model bright and dark sides of personality at work. *Industrial and Organizational Psychology*, *7*(1), 121-126. 10.1111/iops.12119

Wu, J., & LeBreton, J. M. (2011). Reconsidering the dispositional basis of counterproductive work behavior: The role of aberrant personality. *Personnel Psychology*, *64*(3), 593-626.
doi:10.1111/j.1744-6570.2011.01220.x

Table 1

Item pairs

| Item (Item Number) | Constructs |
|--|-------------------|
| I tend to lack remorse (P36) | Psychopathy |
| I am willing to be unethical if I believe it will help me succeed (M21) | Machiavellianism |
| It's fun to see how far you can push people before they get upset (P60) | Psychopathy |
| I like to use clever manipulation to get my way (M38) | Machiavellianism |
| I tend to not be too concerned with morality or the morality of my actions (P38) | Psychopathy |
| I tend to exploit others towards my own end (M47) | Machiavellianism |
| People who are stupid enough to get ripped off usually deserve it (P7) | Psychopathy |
| I am willing to sabotage the efforts of other people if they threaten my own goals (M22) | Machiavellianism |
| I tend to be callous or insensitive (P37) | Psychopathy |
| I believe that lying is necessary to maintain a competitive advantage over others (M24) | Machiavellianism |
| It's true that I can be mean to others (P31) | Psychopathy |
| Most people can be manipulated (M45) | Machiavellianism |
| For me, what's right is whatever I can get away with (P2) | Psychopathy |
| It's wise to keep track of information that you can use against people later (M41) | Machiavellianism |
| *I would be upset if my success came at someone else's expense (P10.R) | Psychopathy (R) |
| People see me as a natural leader (N41) | Narcissism |
| *I'm a soft-hearted person (P47.R) | Psychopathy (R) |
| I think I am a special person (N9) | Narcissism |

Table 1

(continued)

| | |
|--|------------------|
| Rules are made to be broken (P55) | Psychopathy |
| I can make anybody believe anything I want them to (N35) | Narcissism |
| I've often done something dangerous just for the thrill of it (P50) | Psychopathy |
| I really like to be the center of attention (N30) | Narcissism |
| Looking out for myself is my top priority (P8) | Psychopathy |
| I am apt to show off if I get the chance (N20) | Narcissism |
| It is wise to flatter important people (M6) | Machiavellianism |
| *I am an average person (N48.R) | Narcissism (R) |
| Anyone who completely trusts anyone else is asking for trouble (M1) | Machiavellianism |
| I insist upon getting the respect that is due to me (N14) | Narcissism |
| If I show any weakness at work, other people will take advantage of it (M35) | Machiavellianism |
| If I ruled the world, it would be a much better place (N5) | Narcissism |
| There are things you should hide from other people to preserve your reputation (M43) | Machiavellianism |
| I will never be satisfied until I get all that I deserve (N25) | Narcissism |
| Avoid direct conflict with others because they may be useful in the future (M40) | Machiavellianism |
| I tend to want others to admire me (N50) | Narcissism |
| Everybody likes to hear my stories (N23) | Narcissism |
| I hate being the center of attention (N42.R) | Narcissism (R) |

Table 1

(continued)

| | |
|---|------------------|
| I am more capable than other people (N39) | Narcissism |
| I am a born leader (N36) | Narcissism |
| I would cheat if there was a low chance of getting caught (M23) | Machiavellianism |
| People are only motivated by personal gain (M32) | Machiavellianism |
| I tend to manipulate others to get my way (M46) | Machiavellianism |
| It is safest to assume that all people have a vicious streak and it will come out when they are given a chance (M3) | Machiavellianism |
| It bothers me to hurt other peoples' feelings. (P40.R) | Psychopathy (R) |
| *Cheating is not justified because it is unfair to others (P16.R) | Psychopathy (R) |

Note. A * indicates the item was dropped from the final scale because it had a low loading in the confirmatory factor analysis. For the forced-choice version both items in the pair were dropped, and for the single-stimulus version, only the starred item was dropped.

Table 2

Fit Indices for CFAs

| | χ^2 | <i>df</i> | CFI | TLI | SRMR | RMSEA |
|------------------|------------|-----------|------|------|------|-------|
| Machiavellianism | 1239.09*** | 104 | 0.85 | 0.83 | 0.08 | 0.1 |
| Narcissism | 1018.48*** | 65 | 0.78 | 0.74 | 0.07 | 0.12 |
| Psychopathy | 282.37*** | 44 | 0.95 | 0.94 | 0.04 | 0.07 |

Note. Items with low factor loadings were dropped for the narcissism and psychopathy scales.

Table 3

Narcissism model fit for the Graded Response Model

| Item | Honest | | | Faking | | |
|-------|-------------|-----|------|-------------|-----|------|
| | S- χ^2 | df | p | S- χ^2 | df | p |
| N41 | 105.95 | 95 | 0.21 | 574.42 | 85 | 0.00 |
| N9 | 103.60 | 96 | 0.28 | 100.31 | 82 | 0.08 |
| N35 | 107.78 | 99 | 0.26 | 484.42 | 86 | 0.00 |
| N30 | 95.73 | 88 | 0.27 | 102.12 | 78 | 0.03 |
| N20 | 102.08 | 97 | 0.34 | 320.58 | 82 | 0.00 |
| N48.R | 106.47 | 102 | 0.36 | 191.33 | 89 | 0.00 |
| N14 | 130.78 | 114 | 0.13 | 145.40 | 100 | 0.00 |
| N5 | 94.69 | 103 | 0.71 | 261.19 | 91 | 0.00 |
| N25 | 119.11 | 101 | 0.11 | 156.79 | 88 | 0.00 |
| N50 | 95.99 | 100 | 0.59 | 148.47 | 82 | 0.00 |
| N23 | 201.38 | 118 | 0.00 | 271.16 | 106 | 0.00 |
| N42.R | 99.86 | 102 | 0.54 | 256.18 | 84 | 0.00 |
| N39 | 90.51 | 94 | 0.58 | 552.90 | 85 | 0.00 |
| N36 | 105.95 | 95 | 0.21 | 574.42 | 85 | 0.00 |

Note. Item number match the number in parenthesis in Table 1

Table 4

Psychopathy model fit for the Graded Response Model

| Item | Honest | | | Faking | | |
|-------|-------------|-----|------|-------------|-----|------|
| | S- χ^2 | df | p | S- χ^2 | df | p |
| P36 | 58.05 | 63 | 0.65 | 106.54 | 65 | 0.00 |
| P60 | 96.28 | 73 | 0.04 | 104.06 | 78 | 0.03 |
| P38 | 110.00 | 74 | 0.00 | 125.72 | 78 | 0.00 |
| P7 | 101.44 | 88 | 0.15 | 167.76 | 88 | 0.00 |
| P37 | 73.74 | 74 | 0.49 | 114.68 | 79 | 0.01 |
| P31 | 124.61 | 91 | 0.01 | 494.14 | 88 | 0.00 |
| P2 | 66.98 | 72 | 0.65 | 107.30 | 77 | 0.01 |
| P10.R | 70.99 | 89 | 0.92 | 184.29 | 94 | 0.00 |
| P47.R | 151.05 | 99 | 0.00 | 207.65 | 101 | 0.00 |
| P55 | 111.96 | 99 | 0.18 | 204.45 | 104 | 0.00 |
| P50 | 221.64 | 101 | 0.00 | 381.93 | 104 | 0.00 |
| P8 | 58.05 | 63 | 0.65 | 106.54 | 65 | 0.00 |
| P40.R | 96.28 | 73 | 0.04 | 104.06 | 78 | 0.03 |
| P16.R | 110.00 | 74 | 0.00 | 125.72 | 78 | 0.00 |

Note. Item number match the number in parenthesis in Table 1

Table 5

Machiavellianism model fit for the Graded Response Model

| Item | Honest | | | Faking | | |
|------|-------------|-----|------|-------------|-----|------|
| | S- χ^2 | df | p | S- χ^2 | df | p |
| M21 | 88.67 | 78 | 0.19 | 184.85 | 87 | 0.00 |
| M38 | 88.49 | 89 | 0.50 | 177.04 | 95 | 0.00 |
| M47 | 109.82 | 76 | 0.01 | 153.64 | 80 | 0.00 |
| M22 | 91.44 | 82 | 0.22 | 128.85 | 87 | 0.00 |
| M24 | 90.12 | 83 | 0.28 | 134.98 | 89 | 0.00 |
| M45 | 114.64 | 126 | 0.76 | 263.86 | 132 | 0.00 |
| M41 | 116.36 | 105 | 0.21 | 166.11 | 116 | 0.00 |
| M6 | 118.17 | 132 | 0.80 | 515.08 | 135 | 0.00 |
| M1 | 134.30 | 134 | 0.48 | 185.31 | 143 | 0.01 |
| M35 | 141.24 | 142 | 0.50 | 142.21 | 150 | 0.66 |
| M43 | 149.28 | 134 | 0.17 | 287.40 | 144 | 0.00 |
| M40 | 162.63 | 140 | 0.09 | 409.28 | 147 | 0.00 |
| M23 | 100.95 | 97 | 0.37 | 429.89 | 105 | 0.00 |
| M32 | 132.71 | 133 | 0.49 | 193.05 | 138 | 0.00 |
| M46 | 139.38 | 77 | 0.00 | 121.41 | 81 | 0.00 |

Note. Item number match the number in parenthesis in Table 1

Table 6

Means, standard deviations, and correlations of faking Dark Triad scores

| Variable | <i>M</i> | <i>SD</i> | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---------------|----------|-----------|--------|--------|--------|--------|--------|--------|--------|-------|-------|
| 1. CWB | 1.25 | 0.36 | | | | | | | | | |
| 2. MUPP-2PL-N | 0.22 | 0.60 | -.07 | | | | | | | | |
| 3. MUPP-2PL-M | -0.51 | 0.50 | .16** | -.03 | | | | | | | |
| 4. MUPP-2PL-P | -0.59 | 0.48 | .23** | -.16** | .50** | | | | | | |
| 5. TIRT-N | 0.87 | 0.67 | -.16** | .89** | -.26** | -.56** | | | | | |
| 6. TIRT-M | -0.63 | 0.66 | .19** | -.41** | .89** | .61** | -.63** | | | | |
| 7. TIRT-P | -0.77 | 0.65 | .22** | -.68** | .47** | .80** | -.93** | .77** | | | |
| 8. GRM-N | 0.90 | 1.06 | -.02 | .44** | .03 | -.14** | .43** | -.16** | -.35** | | |
| 9. GRM-M | -0.42 | 1.19 | .15** | -.10** | .34** | .15** | -.16** | .34** | .21** | .38** | |
| 10. GRM-P | -0.50 | 1.06 | .20** | -.14** | .29** | .31** | -.26** | .34** | .33** | .22** | .87** |

Note. *M* = mean, *SD* = standard deviation, CWB = counterproductive work behaviors, GRM = graded response model, MUPP-2PL = 2-parameter logistic version of the multi-unidimensional pairwise preference model, TIRT = Thurstonian Item Response Theory Model, -N = narcissism, -M = Machiavellianism, -P = psychopathy. * indicates $p < .05$. ** indicates $p < .01$

Table 7

Means, standard deviations, and correlations of honest DT scores

| Variable | <i>M</i> | <i>SD</i> | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---------------|----------|-----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1. CWB | 1.25 | 0.36 | | | | | | | | | |
| 2. MUPP-2PL-N | -0.50 | 0.59 | .09* | | | | | | | | |
| 3. MUPP-2PL-M | -0.33 | 0.50 | .20** | .25** | | | | | | | |
| 4. MUPP-2PL-P | -0.34 | 0.49 | .21** | .15** | .52** | | | | | | |
| 5. TIRT-N | -0.06 | 0.68 | .10** | .99** | .27** | .15** | | | | | |
| 6. TIRT-M | -0.07 | 0.63 | .23** | .37** | .94** | .71** | .39** | | | | |
| 7. TIRT-P | -0.08 | 0.59 | .22** | .16** | .51** | .97** | .17** | .73** | | | |
| 8. GRM-N | -0.13 | 0.91 | .17** | .51** | .16** | .10** | .51** | .20** | .10** | | |
| 9. GRM-M | -0.16 | 0.92 | .32** | .09* | .34** | .20** | .10** | .33** | .21** | .51** | |
| 10. GRM-P | -0.19 | 0.86 | .35** | .09* | .24** | .34** | .09* | .29** | .34** | .47** | .84** |

Note. *M* = mean, *SD* = standard deviation, CWB = counterproductive work behaviors, GRM = graded response model, MUPP-2PL = 2-parameter logistic version of the multi-unidimensional pairwise preference model, TIRT = Thurstonian Item Response Theory Model, -N = narcissism, -M = Machiavellianism, -P = psychopathy. * indicates $p < .05$. ** indicates $p < .01$.

Table 8

Means, standard deviations, and correlations of HEXACO and CWB scores

| Variable | <i>M</i> | <i>SD</i> | 1 | 2 | 3 | 4 | 5 | 6 |
|---------------------------|----------|-----------|--------|-------|--------|-------|-------|-------|
| 1. CWB | 1.25 | 0.36 | | | | | | |
| 2. Honesty-Humility | 3.36 | 0.70 | -.28** | | | | | |
| 3. Emotionality | 3.25 | 0.68 | -.04 | .03 | | | | |
| 4. Extraversion | 3.15 | 0.72 | -.07 | .10** | -.27** | | | |
| 5. Agreeableness | 3.19 | 0.65 | -.20** | .32** | -.09* | .27** | | |
| 6. Conscientiousness | 3.83 | 0.62 | -.31** | .33** | -.07 | .27** | .22** | |
| 7. Openness to experience | 3.61 | 0.70 | -.12** | .05 | .05 | .13** | .09* | .16** |

Note. *M* =mean, *SD* =standard deviation, CWB = counterproductive work behaviors

* indicates $p < .05$. ** indicates $p < .01$.

Table 9

Comparison of faking and honest scores

| Trait | Honest Mean | Honest <i>SD</i> | Faking Mean | Faking <i>SD</i> | Mean Difference | Cohen's <i>d</i> | <i>t</i> | <i>r</i> |
|------------|-------------|------------------|-------------|------------------|-----------------|------------------|----------|----------|
| MUPP-2PL-N | -0.5 | 0.59 | 0.22 | 0.6 | 0.72 | 1.22 | 24.66** | 0.14** |
| MUPP-2PL-M | -0.33 | 0.5 | -0.51 | 0.5 | -0.18 | -0.36 | -7.87** | 0.26** |
| MUPP-2PL-P | -0.34 | 0.49 | -0.59 | 0.48 | -0.24 | -0.5 | -11.95** | 0.38** |
| TIRT-N | -0.06 | 0.68 | 0.87 | 0.67 | 0.92 | 1.35 | 26.86** | 0.1** |
| TIRT-M | -0.07 | 0.63 | -0.63 | 0.66 | -0.56 | -0.89 | -18.41** | 0.22** |
| TIRT-P | -0.08 | 0.59 | -0.77 | 0.65 | -0.69 | -1.17 | -24.48** | 0.28** |
| GRM-N | -0.13 | 0.91 | 0.9 | 1.06 | 1.02 | 1.13 | 21.75** | 0.2** |
| GRM-M | -0.16 | 0.92 | -0.42 | 1.19 | -0.26 | -0.28 | -5.12** | 0.2** |
| GRM-P | -0.19 | 0.86 | -0.5 | 1.06 | -0.31 | -0.36 | -7.19** | 0.32** |

Note. N = 705, *SD* = standard deviation, *r* = correlation between honest and faking scores, *t* = test statistic from t-test for dependent means, GRM = graded response model, MUPP-2PL = 2-parameter logistic version of the multi-unidimensional pairwise preference model, TIRT = Thurstonian Item Response Theory Model, -N = narcissism, -M = Machiavellianism, -P = psychopathy, Cohen's *d* was calculated by subtracting the honest mean from the faking mean and dividing by the honest standard deviation
* indicates $p < .05$. ** indicates $p < .01$.

Table 10

Hypothesis 1 results

| Trait | MFC mean | GRM mean | <i>t</i> | <i>p</i> |
|------------------|----------|----------|----------|----------|
| MUPP-2PL | | | | |
| Narcissism | 1.22 | 1.127 | 2.126 | 0.034 |
| Machiavellianism | -0.361 | -0.285 | -1.509 | 0.132 |
| Psychopathy | -0.497 | -0.359 | -2.757 | 0.006 |
| TIRT | | | | |
| Narcissism | 1.348 | 1.127 | 5.05 | < .001 |
| Machiavellianism | -0.893 | -0.285 | -11.794 | < .001 |
| Psychopathy | -1.167 | -0.359 | -15.739 | < .001 |

Note. Means for MFC and GRM (single-stimulus) scores are the averages of the standardized scores. Scores were standardized by subtracting the respective honest mean and dividing by the honest standard deviation

Table 11

Hypothesis 2 results

| Trait | Honest | Faking | CI of Difference |
|------------------|--------|---------|------------------|
| MUPP-2PL | | | |
| Narcissism | 0.09* | -0.07 | (0.06, 0.25) |
| Machiavellianism | 0.20** | 0.16** | (-0.08, 0.14) |
| Psychopathy | 0.21** | 0.23** | (-0.15, 0.10) |
| TIRT | | | |
| Narcissism | 0.10** | -0.16** | (0.14, 0.38) |
| Machiavellianism | 0.23** | 0.19** | (-0.09, 0.16) |
| Psychopathy | 0.22** | 0.22** | (-0.13, 0.13) |
| GRM | | | |
| Narcissism | 0.17** | -0.02 | (0.12, 0.27) |
| Machiavellianism | 0.32** | 0.15** | (0.08, 0.26) |
| Psychopathy | 0.35** | 0.20** | (0.05, 0.25) |

Note. Honest = correlation between honest responses and counterproductive work behaviors, Faking = correlation between faking responses and counterproductive work behaviors, CI of Difference = Confidence interval of the difference between Honest and Faking correlations

Table 12

Hypothesis 3 results

| Trait | MFC | SS | CI of Difference |
|------------------|-------|-------|------------------|
| MUPP-2PL | | | |
| Narcissism | -0.07 | -0.02 | (-0.11, 0.03) |
| Machiavellianism | 0.16 | 0.15 | (-0.09, 0.12) |
| Psychopathy | 0.23 | 0.20 | (-0.10, 0.15) |
| TIRT | | | |
| Narcissism | -0.16 | -0.02 | (-0.23, -0.05) |
| Machiavellianism | 0.19 | 0.15 | (-0.07, 0.15) |
| Psychopathy | 0.22 | 0.20 | (-0.10, 0.14) |

Note. MFC= correlation between multidimensional forced choice assessment score and counterproductive work behaviors, SS = correlation between single-stimulus assessment score and counterproductive work behaviors, CI of Difference = Confidence interval of the difference between MFC and SS correlations

Table 13

Research Question results

| | Step 1 | Step 2: MUPP-2PL | Step 2: TIRT | Step 2: GRM |
|-------------------|----------|---------------------|-----------------|----------------|
| (Intercept) | 2.38*** | 2.37*** | 2.32*** | 2.02*** |
| Honest-Humility | -0.08*** | -0.06** | -0.06** | -0.05* |
| Emotionality | -0.02 | -0.03 | -0.03 | 0.01 |
| Extraversion | 0.02 | 0 | 0 | 0.01 |
| Agreeableness | -0.06** | -0.06** | -0.06** | -0.03 |
| Conscientiousness | -0.14*** | -0.12*** | -0.12*** | -0.11*** |
| Openness | -0.03 | -0.03 | -0.03 | -0.03 |
| Narcissism | | 0.02 | 0.02 | 0.01 |
| Machiavellianism | | 0.03 | 0.03 | -0.01 |
| Psychopathy | | 0.07* | 0.06 | 0.09** |
| R ² | 0.15 | 0.17 | 0.17 | 0.17 |
| ΔR^2 | | 0.02* | 0.02* | 0.03*** |

Note. Values are regression slopes. N = 705, GRM = graded response model, MUPP-2PL = 2-parameter logistic version of the multi-dimensional pairwise preference model, TIRT = Thurstonian Item Response Theory Model, * indicates $p < .05$. ** indicates $p < .01$. *** indicates $p < .001$.

Table 14

Frequency of choosing the narcissism option in the MFC assessment in the faking condition

| | Psychopathy | Narcissism |
|------------|------------------|------------|
| P10.R, N41 | 161 | 544 |
| P47.R, N9 | 327 | 378 |
| P55, N35 | 155 | 550 |
| P50, N30 | 199 | 506 |
| P8, N20 | 290 | 415 |
| | Machiavellianism | Narcissism |
| M6, N48.R | 453 | 252 |
| M1, N14 | 240 | 465 |
| M35, N5 | 276 | 429 |
| M43, N25 | 258 | 447 |
| M40, N50 | 275 | 430 |

Note. Item number match the number in parenthesis in Table 1

Table 15

Levene's test for equality of variances

| Trait | Honest Variance | Faking Variance | F | p |
|------------|-----------------|-----------------|-------|--------|
| MUPP-2PL-N | 0.35 | 0.36 | 1.14 | 0.29 |
| MUPP-2PL-M | 0.25 | 0.25 | 0 | 0.96 |
| MUPP-2PL-P | 0.24 | 0.23 | 0.94 | 0.33 |
| TIRT-N | 0.47 | 0.45 | 0.67 | 0.41 |
| TIRT-M | 0.39 | 0.44 | 0.6 | 0.44 |
| TIRT-P | 0.35 | 0.43 | 4.81 | 0.03 |
| GRM-N | 0.82 | 1.13 | 18.55 | < .001 |
| GRM-M | 0.84 | 1.43 | 57.17 | < .001 |
| GRM-P | 0.73 | 1.13 | 64.22 | < .001 |

Note. N = 705, F = test statistic from Levene's test, GRM = graded response model, MUPP-2PL = 2-parameter logistic version of the multi-unidimensional pairwise preference model, TIRT = Thurstonian Item Response Theory Model, -N = narcissism, -M = Machiavellianism, -P = psychopathy

Table 16

Regression slopes for HEXACO predicting faking scores

| Predictor | Narcissism <i>b</i> 's | | | Machiavellianism <i>b</i> 's | | | Psychopathy <i>b</i> 's | | |
|------------------------|------------------------|----------|---------|------------------------------|---------|---------|-------------------------|----------|---------|
| | MUPP-2PL | TIRT | GRM | MUPP-2PL | TIRT | GRM | MUPP-2PL | TIRT | GRM |
| Intercept | 0.29 | 0.32 | 0.21 | -0.09 | 0 | -0.5 | 0.12 | -0.06 | -0.61 |
| Honest Score | 0.21*** | 0.17*** | 0.29*** | 0.26*** | 0.22*** | 0.33*** | 0.33*** | 0.27*** | 0.46*** |
| Honesty-Humility | 0 | 0.01 | 0.01 | 0.01 | 0.02 | 0.2* | 0.01 | 0.03 | 0.24*** |
| Emotionality | 0 | 0.02 | 0.05 | -0.04 | -0.07 | -0.03 | -0.02 | -0.04 | 0.05 |
| Extraversion | -0.15*** | -0.19*** | -0.14* | -0.01 | 0.04 | 0.14 | 0.03 | 0.09** | 0.1 |
| Agreeableness | 0.04 | 0.05 | 0.07 | 0.05 | 0.03 | 0.05 | -0.01 | -0.04 | 0.09 |
| Conscientiousness | 0.02 | 0.1* | 0.07 | -0.05 | -0.08 | -0.16 | -0.07* | -0.08* | -0.21** |
| Openness to Experience | 0.08* | 0.14*** | 0.13* | -0.05 | -0.1** | -0.12 | -0.09*** | -0.14*** | -0.17** |

Note. GRM = graded response model, MUPP-2PL = 2-parameter logistic version of the multi-unidimensional pairwise preference model, TIRT = Thurstonian Item Response Theory Model

APPENDIX

Appendix: Scale Development

The construction of the assessment was based on the sequential approach for creating pairwise preference items described by Stark and colleagues (2005). This process consists of: 1) developing a large number of items for each construct, 2a) administering the items to judges to rate desirability, 2b) administering the items to a group of respondents, 3) estimating item parameters, 4) pairing statements with similar desirability, 5) administering paired statements to respondents, and 6) scoring the assessment. The first four steps were completed in the pilot study, and steps five and six were part of the main study.

I made several modifications to these steps in order to integrate this approach with classic recommendations for scale development (Hinkin, 1998) and to take advantage of the benefits of newer IRT models (i.e., TIRT & MUPP-2PL). First, a content validity assessment was incorporated into the item development. Second, the data collection and item parameter estimation (steps 2b and 3) are necessary in order to score the assessment with the MUPP model, but these steps are not necessary for the TIRT and MUPP-2PL. Rather than omit this data collection, I collected initial single-stimulus responses to determine the factor structure of the DT traits. The dimensionality of the items needs to be known because mapping items to a relevant dimension is an input required by all of the IRT models used to score forced-choice assessments. Narcissism, psychopathy and Machiavellianism have each been defined as multidimensional constructs (Wu & LeBreton, 2011). Before a forced-choice assessment could be created, it was necessary to determine whether it would be more appropriate to consider each of the DT traits unidimensional or have multiple dimensions for each trait. Thus, I collected data (step 2b) but rather than determining item parameters I examined the dimensionality of the DT traits.

Step 1: Item Generation

The first step in constructing pairwise preference items is developing a large pool of items. Because many scales already exist for measuring dark traits, I reviewed and revised existing items rather than generating new items. A description of the seven scales used to create the item pool can be found in Appendix Table 1. I examined each item in these scales and removed 12 items that asked directly about behaviors rather than a belief or attitude. Whether the item measures behaviors versus beliefs/attitudes is an important distinction because behavioral items may overlap with CWB items and other relevant criteria. For example, “stole money from my parents” is directly asking about theft, so including this and similar items could artificially inflate the relationship between psychopathy and CWBs.

Content validity. Because I used existing DT items that were all created by different authors with slightly different definitions of the traits, it was especially important to ensure that all of the items I ultimately chose accurately reflected the intended construct. In order to determine appropriateness of the items, I asked subject matter experts (SMEs) with graduate training in industrial-organizational psychology and psychometrics to review the item pool. They were all given the same definition of each of the DT traits and asked to rate the relevance of each item to its respective construct definition as “essential,” “useful but not essential,” or “not necessary.” Seven doctoral students and one professional with a PhD in industrial-organizational psychology completed the ratings. I calculated Lawshe’s (1975) content validity ratio (CVR) for each item to determine which items were considered essential to the construct. All items with a CVR less than .25, which indicated less than three out of eight SMEs considered the item essential, were removed from the item pool before the desirability ratings. These items were not removed from the data collection in step 2b because that data collection occurred at the same time as the content validity ratings. After summarizing all of the information about the items

available after step three (desirability, factor loading), I was able to use a more stringent cut off of a CVR greater than or equal to .25 (at least five out of eight SMEs considered the item essential) for inclusion in the final scale.

In addition to relevance, SMEs were also asked to indicate whether each item was self- or other-oriented. In light of the finding that items in one specific scale (the Dirty Dozen; Jonason & Webster, 2010) artificially inflate the relationship between narcissism and Machiavellianism by including items that are “other-oriented” on those scales and items that are “self-oriented” on psychopathy (Muris et al., 2017), an effort was made to include a balance of self- and other-oriented items across the constructs. Each item was categorized as either self- or other-oriented if more than half of the SMEs chose the same category. If the SMEs were evenly split, then I indicated that the item was could not be categorized. All of the final items could be categorized. Although not the most important factor in the scale construction, this information was collected to ensure that no scale was made up of exclusively self-oriented or other-oriented items.

Step 2a: Desirability

Social desirability is a tendency to respond to a measure in a way that the respondent believes is socially acceptable (Edwards, 1953). Stark and colleagues (2011) describe two possible methods for obtaining social desirability ratings that have led to similar results: asking respondents to judge desirability or instructing respondents to fake good. Stark and colleagues (2005) recommended using social desirability ratings provided by a panel of judges, but Stark et al. (2011) caution that this method increases cognitive strain on respondents. Researchers have used the direct social desirability ratings successfully (e.g., Usami, Sakamoto, Naito, & Abe, 2016), and I addressed the additional cognitive load by using a panel of 12 SMEs who had received graduate level psychometrics training. These SMEs understand issues of social

desirability responding, so it should have been easier for them to rate social desirability than people without this training.

SMEs were informed that the purpose of the ratings was to determine the desirability of each item so that items of similar desirability could be paired to create an assessment that was resistant to faking. Then they were instructed to rate each of the 130 items with a CVR greater than or equal to $-.25$ on a scale from 1 (*not at all socially desirable*) to 5 (*very socially desirable*). I calculated the mean and average deviation of the social desirability ratings from the 12 raters. As expected, the mean social desirability ratings were lower for psychopathy and Machiavellianism than for narcissism, and the social desirability for the reverse-coded items tended to be much higher than other items (see Appendix Table 2). Average deviation was used to consider the stability of the mean estimates to inform pairing items in step four.

Step 2b: Data Collection

I administered all 183 items to a large pool of respondents in single-stimulus format. These responses were used to determine the factor structure of each DT trait.

Participants. Participants were 607 workers from Amazon's Mechanical Turk (MTurk). They received \$0.50 for participation. Inclusion criteria required participants to be residents of the United States and work at least 30 hours per week. Approximately 48% of the sample were men. The participants had a mean age of 35.80 ($SD = 10.79$). Seventy-eight percent of the sample identified as White, 9% Black or African American, 7% Asian and/or Asian American, and 6% other races. Seven percent of the sample reported their ethnicity as Hispanic or Latino. Participants worked an average of 41.50 ($SD = 5.96$) hours per week.

Procedure. Participants read a brief summary of the study on MTurk and were then directed to the questionnaire on Qualtrics. They completed demographic information followed by

the dark personality items. When they had finished these items, they were given a randomly generated completion code and paid through MTurk.

Measures. The dark personality comprised of the same 183 items reviewed by SMEs for content validation (see Appendix Table 1). Items intended to measure psychopathy, narcissism, and Machiavellianism were combined in a random order before data collection and presented in that same order to participants. All items were rated on a 5-point Likert-type scale (1 = *strongly disagree* to 5 = *strongly agree*).

Step 3: Dimensionality

I used an Exploratory Factor Analysis (EFA) with an oblimin rotation and principle axis factoring to determine the factor structure of the dark personality items. I conducted a separate EFA for narcissism, Machiavellianism, and psychopathy to determine whether each of the DT traits should be considered unidimensional or multidimensional. Only the 85 items that had a CVR greater than or equal to .25 (or 5 out of 8 SMEs defining as essential) were included in the EFAs. The items with a CVR less than .25 were excluded because these items did not accurately reflect their respective constructs. In order to be considered unidimensional for IRT analyses, the first factor should account for more than 20% of the variance (Reckase, 1979). This criterion was met for narcissism (32%), Machiavellianism (28%), and psychopathy (32%). However, it was important to further investigate to determine whether it would be more appropriate to measure each DT trait as multidimensional. There are several methods that can be used to determine the number of factors retained in an EFA. I based my determination upon examination of the scree plot (see Figure 1), parallel analysis (Horn, 1965), percent variance explained by each additional factor, and interpretability of the factors (Fabrigar, Wegener, MacCallum, & Strahan, 1999).

Narcissism. The scree plot for narcissism (see Figure 1) depicts a large change in slope after the first eigenvalue (9.15), which would support a one-factor solution. However, the interpretation of a scree plot is fairly subjective, and it could be argued that there is a significant change in slope, or elbow, after the second eigenvalue (1.44). A more objective method, the parallel analysis, suggested a seven-factor solution. However, this is likely overfactoring because seven factors are not interpretable. In addition, although Kaiser's criterion (Kaiser, 1960) can also lead to overfactoring, it can be used as an upper bound to the number of factors retained (Hayton, Allen, & Scarpello, 2004; Horn, 1965). There were only two factors with an eigenvalue greater than one, so there should be no more than two factors retained. Based on examination of the items with an absolute value of the factor loading of greater than or equal to 0.35 on the second factor, it could be interpreted as attention seeking. For example, the highest loading item was "I like to be the center of attention." However, the addition of a second factor only explained a rather insignificant 5.5% of the variance. Based on the scree plot and the small amount of additional variance explained by adding factors, I concluded that a one-factor solution was appropriate for narcissism.

Machiavellianism. The scree plot for Machiavellianism (Figure 1) also showed a large change in slope after the first eigenvalue (9.01). However, there was a smaller elbow after the third eigenvalue, and both the second (1.74) and third (1.41) eigenvalues are greater than one. Similar to the results for narcissism, the parallel analysis for Machiavellianism suggested a large number of factors (five) that were difficult to interpret. Examining additional variance explained by adding and second (6.3%) and third (5.1%) factor provided more useful information for choosing between a one- or three-factor solution. The addition of each factor accounted for less than 10% of the variance. Finally, I examined the interpretability of the factors. The first factor

was a combination of abstract morality and manipulation (e.g. “I am willing to be unethical if it will help me succeed” and “I tend to manipulate others to get my way”). The second factor appeared to be measuring distrust (e.g. “It’s not wise to tell your secrets” and “Anyone who completely trusts anyone else is asking for trouble”). The third factor was made up of only reverse-coded items (e.g. “Most people are basically good and kind”). A two-factor solution combined the first two factors and simply separated the reverse-coded items. Because the purpose of this analysis was to choose factors to be included as separate dimensions of dark personality, separating reverse-coded items into another factor is not beneficial. The reverse-coded factor does not represent a separate part of dark personality but rather an inversely related part of bright side personality. In addition, pairing items from that factor with items from dark dimensions based on social desirability would be impossible. Therefore, based on the small additional variance explained and the lack of practical utility for the present purpose, I chose a one-factor solution for Machiavellianism.

Psychopathy. The scree plot for psychopathy could indicate either a one- or two-factor solution (see Figure 1), and both the first (7.74) and second (1.78) eigenvalues are greater than one. The parallel analysis suggests a four-factor solution, which was not interpretable. The second factor only explained an additional 8.3% of the variance, and the second factor was comprised of all reverse-coded items. The reverse-coded factor is not a useful dimension to build a forced-choice test of dark personality, and it does not explain much additional variance, so a one-factor solution is appropriate for psychopathy.

Step 4: Pairing Items

Finally, items from different facets were paired based on similarity of social desirability ratings. I chose to use item pairs rather than larger blocks of items for two reasons. First,

increasing the size of the blocks increases how long it takes respondents to complete the measure (Sass et al., 2018), which is likely due to higher cognitive complexity. In a selection context, personality measures are used to provide incremental validity over cognitive ability, so saturation with cognitive ability is both redundant and likely to increase adverse impact. Second, the MUPP-2PL offers important advantages over other models (e.g., simultaneous estimation of item and person parameters, freedom in assessment construction) and is only valid for pairs.

Only the items that had a CVR greater than or equal to .25 (or 5 out of 8 SMEs rated essential) and the absolute value of the factor loading was greater than or equal to .35 on the relevant one-factor EFA were considered. Items were sorted by mean social desirability and paired with items on other dimensions with similar mean social desirability values. The mean was less informative for items with a high average deviation in social desirability ratings (> 0.80 ; Burke & Dunlap, 2002), so the median of the social desirability ratings was also considered. As I paired items I reviewed them to ensure that one item was not clearly more desirable. If one item in a pair seemed “better,” then it was paired with an item that had a slightly higher, but still similar mean social desirability.

An additional consideration was that the TIRT model has better parameter recovery when pairs are keyed in opposite directions (i.e., one item is reverse coded and the other is not; Brown & Maydeu-Olivares, 2011), so an effort was made to match these opposite polarity items. However, opposite polarity blocks are less robust to faking, so only opposite polarity items with similar desirability scores were paired, as the purpose is to develop a faking-resistant assessment. I created three multidimensional opposite polarity items.

In addition to multidimensional pairs, I created one unidimensional pair of psychopathy items and two unidimensional pairs of narcissism and Machiavellianism items in order to anchor

the scale. Items chosen for unidimensional pairs should ideally have different factor loadings so that they provide information needed to determine the scale origin (Brown, 2016), but within the constraints of pairing items with similar social desirability, factor loadings were not largely different for four of the five unidimensional pairs. One of the narcissism unidimensional pairs contained a reverse-coded and normal item with same social desirability, so this item had the largest difference in factor loadings (1.04). The mean of absolute value of differences in factor loadings for the other four items was 0.16.

After the first draft of pairs was complete, I checked to be sure no dimension was exclusively comprised of self-oriented or other-oriented items. Finally, I reviewed the definition of each of the DT constructs that I gave to the SMEs for relevance ratings to determine whether every part of the definition was covered by at least one item. A few items were switched to ensure content coverage for each trait. Not all aspects of the definition were equally represented, but every part of the definition was addressed by at least one item. Again, no dimension was exclusively comprised of self-oriented or other-oriented items. After this final revision, the mean of the absolute value of the differences between social desirability ratings for items pairs was 0.20 ($SD = .18$). The final set of 22 item pairs can be found in Table 1 of the manuscript.

Table 1

Summary of item pool

| Scale | Number of Items | Number of Items Chosen | Citation | Construct(s) |
|--|------------------------|-------------------------------|---|---|
| Narcissistic Personality Inventory (NPI) | 40 | 10 | Raskin & Hall (1979) | Narcissism |
| MACH-IV | 20 | 3 | Christie & Geis (1970) | Machiavellianism |
| Levenson Self-Report Psychopathy Scale | 26 | 5 | Levenson et al. (1995) | Psychopathy |
| Self-report psychopathy scale (SRP-III) | 34 | 5 | Mahmut, Menictas, Stevenson, & Homewood (2011)* | Psychopathy |
| Machiavellian Personality Scale (MPS) | 16 | 6 | Dahling et al. (2009) | Machiavellianism |
| Short Dark Triad (SD3) | 27 | 9** | Jones & Paulhus (2014) | Narcissism, Machiavellianism, Psychopathy |
| Dirty Dozen (DD) | 12 | 6** | Jonason & Webster (2010) | Narcissism, Machiavellianism, Psychopathy |

Note. *Paulus, Neumann, & Hare (2009) created a 40-item version, but the 34-item version updated by Mahmut, Menictas, Stevenson, & Homewood (2011) was used.

**The 9 SD3 items chosen included 5 Machiavellianism items, 3 narcissism items, and 1 psychopathy item. The 6 DD items chosen included 2 Machiavellianism items, 1 narcissism item, and 3 psychopathy items.

Table 2

Summary of social desirability ratings

| All Rated | | | |
|-------------------|----|-------------------|------|
| Construct | N | Mean Desirability | SD |
| Psychopathy | | | |
| Not Reversed | 26 | 1.59 | 0.57 |
| Reversed | 17 | 4.15 | 0.58 |
| Narcissism | | | |
| Not Reversed | 35 | 2.96 | 0.73 |
| Reversed | 11 | 3.16 | 0.78 |
| Machiavellianism | | | |
| Not Reversed | 30 | 1.77 | 0.56 |
| Reversed | 11 | 4.14 | 0.48 |
| Considered | | | |
| Construct | N | Mean Desirability | SD |
| Psychopathy | | | |
| Not Reversed | 14 | 1.56 | 0.46 |
| Reversed | 10 | 4.14 | 0.72 |
| Narcissism | | | |
| Not Reversed | 24 | 2.94 | 0.65 |
| Reversed | 5 | 3.02 | 0.45 |
| Machiavellianism* | | | |
| Not Reversed | 32 | 2.38 | 1.17 |

Note. N = number of items. Mean Desirability = Mean social desirability rating for 12 SMEs. SD = Standard deviation of the Mean Desirability. All Rated = summary of the 130 items that were rated by SMEs. Considered = summary of the 85 items with a CVR \geq .25 and absolute value of the factor loading \geq .35 that were considered for pairing.

*No reverse-coded Machiavellianism items were considered

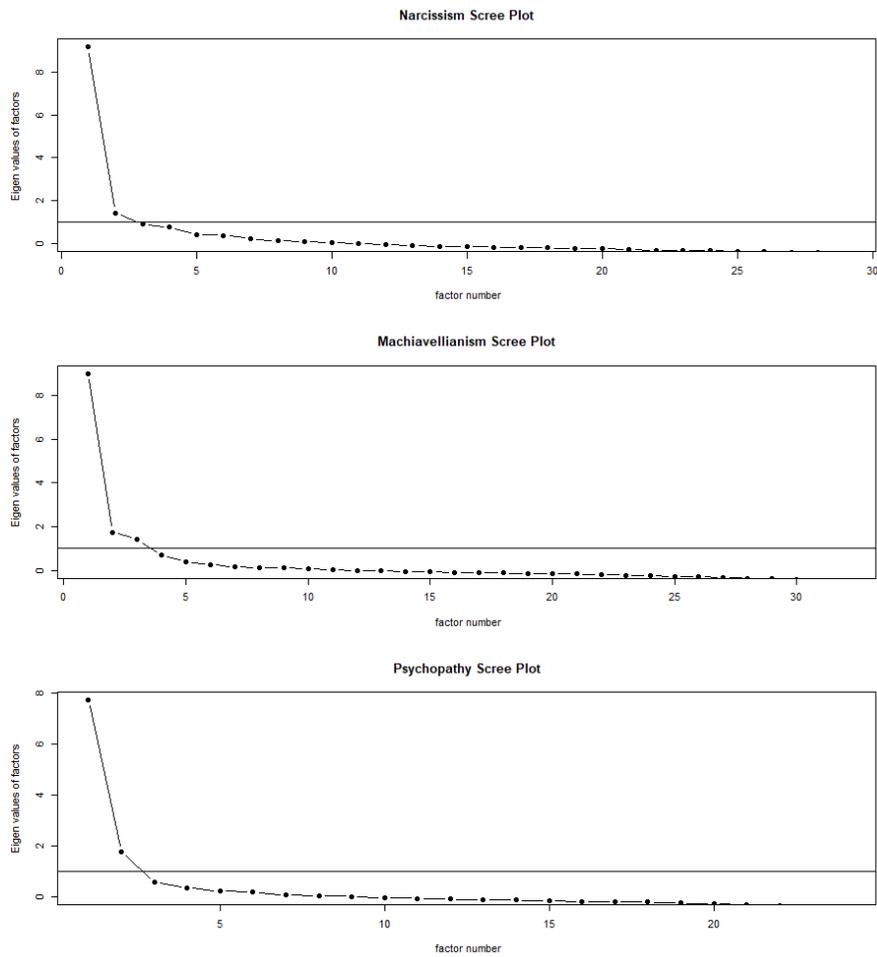


Figure 1. Scree plots for pilot study EFAs. The narcissism, Machiavellianism, and psychopathy scree plots contain 32, 29, and 24 items, respectively.