
AFRL-RH-WP-TR-2020-0106

# DEVELOPMENT OF DARK TETRAD RESEARCH FORMS FOR THE U.S. AIR FORCE

**Fritz Drasgow**
**Oleksandr S. Chernyshenko**
**Stephen Stark**
**Christopher D. Nye**
**Bo Zhang**
**Tianjun Sun**
**Lingyue Li**

**Drasgow Consulting Group**

**November 2020**

**Interim Report**



**AIR FORCE RESEARCH LABORATORY**
**711TH HUMAN PERFORMANCE WING,**
**AIRMAN SYSTEMS DIRECTORATE,**
**WRIGHT-PATTERSON AIR FORCE BASE, OH 45433**
**AIR FORCE MATERIEL COMMAND**
**UNITED STATES AIR FORCE**

## NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 88th Air Base Wing Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RH-WP-TR- 2020-0106 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//signature//
THOMAS R. CARRETTA
Work Unit Manager
Performance Optimization Branch
Airman Biosciences Division

//signature//
LOGAN A. WILLIAMS, DR-III, PhD
Core Research Area Lead
Performance Optimization Branch
Airman Biosciences Division

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE *(DD-MM-YY)* | 2. REPORT TYPE | 3. DATES COVERED *(From - To)* |
|---|---|---|
| 16-11-20 | Interim | 29 October 2019 – 16 November 2020 |

**4. TITLE AND SUBTITLE**

Development of Dark Tetrad Research Forms for the U.S. Air Force

**5a. CONTRACT NUMBER**
FA8650-14-D-6500, Task Order 0007

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**
62202F

**6. AUTHOR(S)**
Fritz Drasgow, Oleksandr S. Chernyshenko, Stephen Stark, Christopher D. Nye, Bo Zhang, Tianjun Sun, and Lingyue Li

**5d. PROJECT NUMBER**
5329

**5e. TASK NUMBER**
09

**5f. WORK UNIT NUMBER**
H0SA (532909TC)

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Drasgow Consulting Group (DCG)
3508 Highcross Rd.
Urbana, IL, 6180

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
Air Force Materiel Command
Air Force Research Laboratory
711th Human Performance Wing
Airman Systems Directorate
Airman Biosciences Division
Performance Optimization Branch
Wright-Patterson AFB, OH 45433

**10. SPONSORING/MONITORING AGENCY ACRONYM(S)**
711 HPW/RHBC

**11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)**
AFRL-RH-WP-TR- 2020-0106

**12. DISTRIBUTION/AVAILABILITY STATEMENT**
Distribution A: Approved for public release

**13. SUPPLEMENTARY NOTES**
AFRL-2020-0612, cleared 28 December 2020

**14. ABSTRACT**
Two forms were developed to assess the Dark Tetrad traits. Both assess psychopathy, narcissism, Machiavellianism, and sadism as well as four Bright side facets (achievement, even tempered, selflessness, and virtue) of the Big Five personality dimensions. The versions include a traditional single statement format with a 5-point Likert response scale and a two-alternative forced-choice format with multidimensional paired statements. Software for scoring the response formats was also developed. The statements were administered to a sample of USAF Basic Recruits and MTurk workers to assess their psychometric characteristics. After the statements were calibrated, the forms were assembled and administered to samples of MTurk workers and Prolific workers. Reasonably good cross-form correlations were found as well as substantial correlations with alternative measures of the same construct. The single statement versions of the Dark Tetrad traits showed reasonable resistance to faking good.

**15. SUBJECT TERMS**
Dark Tetrad traits; Bright Side personality traits; Likert scale; personality assessment; Big Five model.

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT: | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON (Monitor) |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | Thomas R. Carretta |
| Unclassified | Unclassified | Unclassified | SAR | 35 | 19b. TELEPHONE NUMBER *(Include Area Code)* N/A |

## TABLE OF CONTENTS

## LIST OF TABLES

## PREFACE

The work described in this technical report was performed under the task "Air Force Tailored Adaptive Personality Assessment System (AF TAPAS) Modification – Dark Tetrad," Contract FA8650-14-D6500, Task Order 7, Enhanced Airman Alignment, WU 532909TC (H0SA). We thank John Trent and the Air Force Personnel Center, Strategic Research and Assessment Branch (AFPC/DSYX) for data collection at Lackland AFB.

**SUMMARY**

The U.S Air Force (USAF) requires static forms of personality inventories for personnel studies. Accordingly, the present study developed two forms that assess four Dark side traits (psychopathy, sadism, narcissism, and Machiavellianism) and four of 15 facets of Bright Side personality traits from the Air Force Tailored Adaptive Personality Assessment System (AF TAPAS (achievement, even tempered, selflessness, and virtue). Approximately 50 statements were written to assess high, low, and intermediate levels of each of the four Dark side traits. They were administered to a sample of USAF Basic Recruits and MTurk workers to assess their psychometric characteristics. Albeit with considerable difficulty, the statements were calibrated with the Generalized Graded Unfolding Model. The two forms were then assembled. The single statement form consists of 58 items where each statement is presented individually, and respondents are asked to indicate agreement on a 5-point Likert response scale. The two-alternative forced-choice (2AFC) form presents respondents with a pair of statements and asks them to choose the statement that is "more like me." Each 2AFC item consists of two personality statements that are similar in extremity and desirability, but represent different personality dimensions. Software to provide scoring for each format was also developed. The forms were administered to samples of MTurk and Prolific workers to assess their psychometric characteristics. Reasonably good cross-form correlations were found as well as substantial correlations with alternative measures of the same construct. The single statement versions of the Dark Tetrad traits showed reasonable resistance to faking good.

## 1.0 INTRODUCTION

The Air Force Tailored Adaptive Personality Assessment System (AF TAPAS) is a noncognitive measure used to assess personality factors related to performance in military specialties. AF TAPAS is a "Bright Side" measure: It assesses facets underlying the Big Five personality traits. The purpose of this work was to develop a measure of four "Dark Side" traits, sadism, psychopathy, narcissism, and Machiavellianism, and evaluate its psychometric characteristics.

### 1.1 Background

The AF TAPAS is a DOD-owned, personality assessment measure rooted in the Big Five theory of personality, containing 15 facets designed to assess personality factors related to performance in military specialties. There are three versions of the instrument: a single statement version using the Likert response format, a two-alternative forced choice (2AFC) version where each pair of statements assess the same underlying trait, and a 2AFC version where the two statements assess different traits (Chernyshenko, Drasgow, Stark, & Nye, 2019). This last version, the multidimensional AF TAPAS, builds on the Army's Assessment for Individual Motivation (AIM) and incorporates features that address problems associated with traditional Likert scale measures of personality traits, including faking, limitations of classical test theory (CTT), and test compromise. The multidimensional AF TAPAS, which corresponds to Army TAPAS Version 5, is only one of several DOD-owned versions of TAPAS. Several of the Army's versions contain facets not included on the AF TAPAS forms.

The dark triad in psychology focuses on three personality traits: narcissism, Machiavellianism, and psychopathy (Paulhaus & Williams, 2002). Use of the term "dark" implies that people possessing these traits have malevolent qualities. People scoring high on these traits are more likely to commit crimes, cause social distress, and create severe problems for an organization (e.g., counterproductive work behavior), especially if they are in leadership positions. A factor analysis which included measures of the Big Five showed that among the Big Five personality traits, low agreeableness was the strongest correlate of the Dark Triad, and that neuroticism and a lack of conscientiousness were associated with some of the Dark Triad traits (Jacobwitz & Egan, 2006). Although the three dark triad traits are conceptually distinct, empirical evidence shows them to overlap to some degree: They are associated with a callous-manipulative interpersonal style (Jones & Paulhaus, 2010).

Narcissism is characterized by egotism, grandiosity, pride, and a lack of empathy (Kohut, 1977). It can be defined as a person's tendency to have a grandiose, exaggerated sense of their own self-importance and to be preoccupied with thoughts and fantasies of great success. Grandiosity may lead narcissists to a never-ending quest for ego-reinforcement (Jones & Paulhus, 2014). The cognitive process underlying narcissism seems to be more self-deceptive than those of Machiavellianism: Narcissists appear to believe their inflated images of themselves. This grandiosity promotes a sense of entitlement, which justifies their exploitation of other people.

Machiavellianism, on the other hand, is typified by carefully calculated manipulation and exploitation of others, a cynical disregard for morality, and a focus on deception and self-interest

(Jacobwitz & Egan, 2006). Machiavellianism is more deliberate than narcissism; Machiavellians are strategic and plot carefully. They take care that their manipulations do not harm their reputations. They may build alliances to implement their plans. Notably, they do not care if others are harmed; they have callous disregard for others.

Psychopathy is characterized by continuing antisocial behavior, callousness, impulsivity, remorselessness, and selfishness (Skeem, Polaschek, Patrick, & Lilienfeld, 2011). There is a notable lack of affect (i.e., callousness) and low self-control, which leads to impulsive behavior. In contrast to the carefully laid plots of Machiavellians, psychopaths enact their callousness in the short-term, which leads to recklessness and thrill seeking.

Psychopathy, narcissism, and Machiavellianism have been labelled the "Dark Triad." O'Boyle, Forsyth, Banks, and McDaniel (2012) reviewed studies of the Dark Triad personality traits and meta-analytically examined their relations to job performance and counterproductive work behavior (CWB). The data set consisted of reports of 245 independent samples *(N* = 43,907) from reports published between 1951 and 2011. The authors concluded that reductions in the quality of job performance were consistently associated with increases in Machiavellianism and psychopathy and that CWB was associated with increases in all three Dark Triad components. It should be noted that these relationships were moderated by contextual factors (e.g., authority and culture). Multivariate analyses indicated that the Dark Triad constructs were associated with moderated amounts of the variance in CWB, but not job performance. Further, these three traits were found to be positively related to one another but sufficiently distinctive to warrant theoretical and empirical partitioning.

Some have suggested expanding the Dark Triad to include a fourth facet, sadism (Mededovic & Petrovic, 2015). Mededovic and Petrovic contend that the Dark Triad traits can be located in the space of basic personality traits, especially at the negative pole of Hostility-Humility, and the dimensions of Agreeableness, Conscientiousness, and Emotional Stability. While sadism behaves in a similar manner to the other dark traits, it cannot be reduced to them and does not fit in the normal range of personality (i.e., Big Five).

Sadism represents a combination of different behavioral, cognitive, and interpersonal characteristics related to pleasure in connection with inflicting emotional or physical pain on others (Reidy, Zeichner, & Seibert, 2011) and to control, punish, and humiliate others (Myers, Burket, & Husted, 2006). Several studies have shown that sadism and the Dark Triad traits, primarily psychopathy, have several common characteristics, such as the lack of empathy and readiness for emotional involvement (Kirsch & Becker, 2007), inflicting suffering on others, and a connection with antisocial behavior (Chabrol, Van Leeuwen, Rodgers, & Sejourne, 2009). These have been shown to be related but distinct traits (Mokros, Osterheider, Rucker, & Nitschke, 2011). Reidy et al. (2011) found that sadism, separately from psychopathy, predicted unprovoked aggression in the laboratory context. Sadism has also been shown to predict delinquent behavior for students separately from other Dark Triad traits (Chabrol et al., 2009).

## 1.2 Objective and Overview

The objective of this effort was to expand the content of the AF TAPAS (Chernyshenko et al., 2019) beyond its current Big Five personality constructs to include facets covering the Dark Tetrad constructs of Machiavellianism, narcissism, psychopathy, and sadism.

Two studies were conducted. In the first, large pools of items were written according to the definitions of the Dark Tetrad Constructs. They were then administered to a sample of MTurk workers and a sample of USAF Basic Recruits. Attempts to estimate item response theory (IRT) item parameters from these data sets were met with considerable difficulty, but eventually estimates were obtained from the combined MTurk-AF recruit data set.

In the second study, the IRT item parameter estimates were used to assemble two forms: a single statement version using a five-alternative Likert format and a two alternative forced-choice (2AFC) version. These forms were administered to MTurk and Prolific samples and the resulting data were analyzed. This report summarizes the two studies and highlights key findings.

# 2.0 STUDY 1: STATEMENT DEVELOPMENT AND CALIBRATION

## 2.1 Construct Definitions

The Dark Tetrad traits have been conceptualized in various ways. To write statements describing high, low, and intermediate aspects of these traits, it was necessary to have clear construct definitions. Moreover, to use unidimensional item response theory for their analysis, it was important to circumscribe their characteristics. A review of the literature on the Dark Tetrad led to the following definitions:

**Narcissism** is defined as a person's tendency to have a grandiose, exaggerated sense of their own self-importance and to be preoccupied with thoughts and fantasies of great success. Key adjectives: Inflated self-importance, attention seeking, and lack of humility.

**Machiavellianism** is defined as a person's tendency to be unemotional, and therefore able to detach oneself from conventional morality and hence to deceive and manipulate others. Key adjectives: Manipulative, scheming, calculating, conniving, deceitful, underhanded, cunning, insidious, deceptive, crooked, and duplicitous.

**Psychopathy** is defined as a person's tendency to act out impulsively and be emotionally and interpersonally detached. The core of this dimension is guiltlessness and lovelessness. Key adjectives: loveless, guiltless, aggressive, uncaring, reckless, rebellious, risk-taking, unsentimental, and callous.

**Sadism** is defined as a person's tendency to be cruel and find cruelty pleasurable and exciting. This enjoyment of cruelty can even occur in everyday life such as enjoying films, sports, or video games that include cruel content. Key adjectives: Cruel, cold-hearted, and hurtful.

## 2.2 Statement Creation

After the definitions of the traits were established, a large number of statements were written to reflect high, low, and intermediate levels. For example, high, intermediate, and low narcissism statements were “I deserve the best of everything,” “I deserve respect, but not more than anybody else,” and “I am not better than anyone else.”

An Excel workbook was created with draft statements for each of the traits on separate worksheets. The draft statements were reviewed, duplicates were deleted, statements straying too far from the construct definitions were deleted, and the remaining statements were edited for clarity. Ultimately approximately fifty statements per trait were developed.

## 2.3 Statement Pretesting

Careful inspection of the statement pools led to 52, 54, 50, and 50 statements being selected for pretesting for narcissism, Machiavellianism, psychopathy, and sadism.

These statements were administered to 500 MTurk workers and 747 USAF Basic Recruits. A four-point Likert response scale (1 = 1 Strongly disagree; 2 = Disagree; 3 = Agree; 4 = Strongly Agree) was used. Seven quality control items were included. These items asked about impossible events (e.g., “I can run 2 miles in 2 minutes”). Data from respondents who answered more than one of these items as anything other than Strongly Disagree was deleted. Also deleted were cases where more than half of the items were omitted and cases where endorsement rates exceeded 90% or fell below 10%.

A total of 419 MTurk respondents and 550 USAF Basic Recruits provided usable data.

## 2.4 Statement Item Response Theory Calibration

Initially, the data from the MTurk sample and the USAF Basic Recruits sample were analyzed separately using the GGUM2004 item response theory (IRT) software. It fits the Generalized Graded Unfolding Model (GGUM) using maximum likelihood estimation. GGUM2004 was run separately for each trait and sample.

Unfortunately, substantial convergence problems were encountered. Either the software did not converge or it converged to sets of item parameters that were clearly inappropriate (e.g., virtually all statements identified as intermediate). Several other software programs were tried, but without success.

There are at least two causes of non-convergence: “Bad items” and excessive multidimensionality. Although testing programs carefully edit items (as did we), it is an empirical fact that some items are failures and have poor psychometric properties. When IRT software tries to fit all items onto a common scale, the inclusion of poorly functioning items can lead to non-convergence. The other source of non-convergence, excessive multidimensionality, can cause problems because GGUM is a unidimensional IRT model. When confronted by excessive multidimensionality, it may fail to converge or converge to a local maximum that is far from the global maximum.

As a result of these two considerations, we tried deleting statements that seemed to have poor psychometric properties or might have been too multidimensional. We had limited success and ultimately decided to combine the MTurk and USAF Basic Recruits data sets to increase the sample size for the GGUM analysis. With the combined data set and reduced set of statements GGUM2004 converged to reasonable solutions for all four Dark Tetrad traits.

## 2.5 Forms Assembly

We had initially planned to add the four Dark Tetrad traits to the AF TAPAS. However, this would have created an instrument assessing 19 traits and would have been quite long. Consequently, it was decided to create a separate measure.

In her undergraduate Honors Thesis project, Lingyue Li had developed a 2AFC measure of the Dark Triad. Her measure included the Dark Triad (narcissism, psychopathy, and Machiavellianism) and three Bright Side traits (conscientiousness, agreeableness, and emotional stability). In an MTurk study, she had obtained good results: strong correlations with single statement measures of these traits and substantial resistance to faking. Based on Li's work, it was decided to create 8 dimensional measures with the four Dark Tetrad traits and the four Bright Side traits of achievement, even tempered, selflessness, and virtue.

The definitions of the four Bright Side traits are:

**Achievement**: High scoring individuals are seen as hard working, ambitious, confident, and resourceful.

**Even Tempered**: High scoring individuals tend to be calm and stable. They do not often exhibit anger, hostility, or aggression.

**Selflessness**: High scoring individuals are generous with their time and resources.

**Virtue**: High scoring individuals adhere to standards of honesty, morality, and "good Samaritan" behavior.

### 2.5.1. 56-item Single Statement Form

The single statement (SS) form utilizes the traditional, Likert format, where each statement is presented individually, and respondents are asked to indicate agreement on a 5-point Likert scale (1 = Strongly disagree; 2 = Disagree; 3 = Neither disagree not agree; 4 = Agree; 5 = Strongly agree). The Likert format has been used widely by personality researchers to specify the hierarchical structure of personality traits and to estimate criterion related validities in educational, health, and employment contexts. Example personality measures utilizing the Likert format include the NEO Personality Inventory-Revised (Costa & McCrae, 2008), Trait Self-Description Inventory (TSDI), and Assessment of Background and Life Experiences (ABLE; White & Young, 1998).

This form consists of 56 personality single-statement items that are distributed evenly across the four Dark Tetrad and four Bright Side personality facets; two quality control items were also included in the study reported here. Bright Side statements were selected from the TAPAS research statement pool. Selected statements had large discrimination parameters and described either positive or negative standings on the trait continuum (a.k.a., positively/negatively worded). Negatively worded statements were included because they help to combat the acquiescence bias typically associated with the Likert format (i.e., a tendency to agree with all items in a scale regardless of their content). All statements were then randomly ordered and a standard set of instructions was added.

### 2.5.2. 56-item 2AFC Form

The 2AFC form utilizes the multidimensional pairwise preference (MDPP) item format. In this format, items are comprised of two personality statements that are similar in extremity and desirability, but represent different personality dimensions. A small number of unidimensional pairs were also added to facilitate scoring accuracy and to improve examinee reactions. Specifically, there were a total of eight unidimensional pairs, one pair per dimension. Respondents are asked to choose one statement in each pair that is "more like me." Examples of inventories utilizing the MDPP format ar the Army's TAPAS and the multidimensional AF TAPAS.

The form constructed to assess the Dark Tetrad consists of 56 2AFC items and two quality control items. Each personality facet is assessed by 14-15 statements that were selected from the Dark Tetrad and TAPAS statement pools. MDPP items were constructed by matching statements based on extremity (statement location) and desirability. A statement was used only once.

## 2.6 Scoring Procedures

The two personality research forms were designed for paper-and-pencil test administration. To score each form, an on-line scoring tool was developed. This was necessary because the 2AFC form needs to be scored using item response theory (IRT) methods containing complex mathematical routines. To obtain scores for each of the two forms, examinee item response data must be submitted in an excel format (.cvs); resulting test scores are also outputted in this excel format. We briefly describe each scoring routine below.

### 2.6.1. Scoring the SS Form

The scoring routine for this form is straightforward and utilizes the classical test theory (CTT) approach. Each examinee's item responses must be coded as 1 = "strongly disagree", 2 = "disagree", 3 = "neither agree nor disagree", 4 = "agree", and 5 = "strongly agree"; missing responses must be left as blanks. After receiving examinee responses, the scoring routine first reverse scores negatively worded items. Then, for each personality scale, the routine computes the average score across all endorsed items belonging to that scale; if some of the seven items are not endorsed, they do not affect the computation of that scale average. Finally, the scores are outputted in the .csv format.

### 2.6.2. Scoring the 2AFC Form with the MDPP Model

Scoring of this form is by the Multidimensional Pairwise Preference model (MDPP; Stark, Chernyshenko, & Drasgow, 2005) which specifies the probability of endorsing statement *s* over a statement *t* as

$$P_{(s>t)_i}(\theta_{d_s},\theta_{d_t}) = \frac{P_{st}\{1,0\}}{P_{st}\{1,0\}+P_{st}\{0,1\}} \approx \frac{P_s\{1\}P_t\{0\}}{P_s\{1\}P_t\{0\}+P_s\{0\}P_t\{1\}},$$

where:
$i =$ index for items, consisting of pairs of statements, where $i$ = 1 to I,
$d =$ index for dimensions, where $d$ = 1, …, D,
$s,t =$ indices for the first and second statements, respectively, in an item,
$\theta_{d_s},\theta_{d_t} =$ latent trait values for a respondent on dimensions $d_s$ and $d_t$ respectively,
$P_s\{1\}, P_s\{0\} =$ probability of endorsing/not endorsing statement $s$ at $\theta_{d_s}$,
$P_t\{1\}, P_t\{0\} =$ probability of endorsing/not endorsing statement $t$ at $\theta_{d_t}$,
$P_{st}\{1,0\} =$ joint probability of endorsing statement $s$, and not endorsing statement $t$ at $(\theta_{d_s},\theta_{d_t})$,
$P_{st}\{0,1\} =$ joint probability of not endorsing statement $s$, and endorsing statement $t$ at $(\theta_{d_s},\theta_{d_t})$,
and
$P_{(s>t)_i}(\theta_{d_s},\theta_{d_t}) =$ probability of respondent $j$ preferring statement $s$ to statement $t$ in pairwise preference item $i$.

Note that the probabilities of endorsing/not endorsing a stimulus in a pairwise preference item is computed using the Generalized Graded Unfolding Model (GGUM; Roberts, Donoghue, & Laughlin, 2000), GGUM parameters for the TAPAS research pool were estimated using samples of U.S. military recruits undergoing their basic training and parameters for the Dark Tetrad statements are described above.

The scoring of 2AFC response patterns is accomplished via the Bayes modal estimation approach. For a vector of latent trait scores, $\tilde{\theta} = (\theta_{d'=1}, \theta_{d'=2}, \ldots, \theta_{d'=D})$, $\tilde{\boldsymbol{\theta}} = (\theta_{d'=1},\theta_{d'=2},...,\theta_{d'=D})$, this involves maximizing:

$$L(\tilde{\mathbf{u}},\tilde{\boldsymbol{\theta}}) = \{\prod_{i=1}^{n}[P_{(s>t)_i}]^{u_i}[1-P_{(s>t)_i}]^{1-u_i}\} * f(\tilde{\boldsymbol{\theta}}),$$

where $\tilde{\mathbf{u}}$ represents an examinee's item response pattern, $u_i$ is the dichotomous response to item $i$, $P_{(s>t)_i}$ is the probability of preferring statement $s$ to statement $t$ in item $i$, and $f(\tilde{\boldsymbol{\theta}})$ is a $D$-dimensional prior density function, which, for simplicity, is assumed to be the product of independent normals,

$$\prod_{d'=1}^{D}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-\theta_{d'}^2}{2\sigma^2}\right).$$

Taking the natural log, for convenience, the above equation can be rewritten as:

$$\ln L(\tilde{\mathbf{u}}, \tilde{\boldsymbol{\theta}}) = \sum_{i=1}^{n} [u_i \ln P_{(s>t)_i} + (1-u_i)\ln(1-P_{(s>t)_i})] + \sum_{d'=1}^{D} \left[ \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{\theta_{d'}^2}{2\sigma^2} \right] \quad ,$$

leaving the following set of equations to be solved numerically:

$$\frac{\partial \ln L}{\partial \tilde{\boldsymbol{\theta}}} = \begin{bmatrix} \frac{\partial \ln L}{\partial \theta_{d'=1}} \\ \frac{\partial \ln L}{\partial \theta_{d'=2}} \\ \cdots \\ \frac{\partial \ln L}{\partial \theta_{d'=D}} \end{bmatrix} = 0.$$

In total, for each 56-item response pattern submitted, the scoring routine outputs 8 Bayes modal estimates (one per personality construct). Similar to the single statement form, the scores are saved in the .csv format.

## 3.0 TEST OF THE SS AND 2AFC FORMS

### 3.1 Forms and Administrative Conditions

To evaluate the psychometric characteristics of the instruments, the two forms (SS and 2AFC) were adapted for administration via Qualtrics. For one administrative condition, respondents were directed to "answer honestly" and for the second respondents were instructed to "fake good." Specifically, the answer honestly directions were "*There are no right or wrong answers. Simply describe yourself honestly and accurately. In deciding on an answer, consider your life in general and not only the last few weeks or months.*" The fake good directions were "*Imagine that you are applying for a job that you dreamed about, and now you are taking a personality test for that job application. Please answer the following questions by trying your best to create a good impression of yourself from a standpoint of an employee. You do not need to be honest with the questions, just try to choose any answer you think is the best so you can have a better chance of being hired. It is very important that you respond to the statements to present the best of yourself.*" The forms were administered in two orders: (1) Likert and then 2AFC; and (2) 2AFC and then Likert.

### 3.2 Demographic and Criterion Variables

A fairly extensive set of additional variables were assessed. They included eight demographic variables: Education, Income, Employment Status, Age, Organizational Status, Social Media

Usage, Charity Behavior, and Financial Behavior. The demographic variables were assessed by the following methods:

Education: Respondents were asked to indicate the highest level of education they have completed.

Income: Respondents were asked to indicate their current annual household income before tax in US dollars.

Employment status: Respondents were asked whether they are full-time employed, part-time employed, self-employed, a student, retired, or unemployed.

Age: Respondents were asked to report their age.

Organizational status: Respondents were asked to indicate, for their current job, whether they have the authority to hire people, fire people, supervise people, create budgets for the organizational, or make strategic decisions for the organization.

Social media usage: Respondents were asked to indicate the extent to which they spend time on Facebook, Twitter, Instagram, Snapchat, TikTok, Reddit, Tumblr, Pinterest, Quora, LinkedIn, YouTube, and Tinder.

Charity behavior: Respondents were asked to indicate during the past year whether they had donated money, donated blood, donated clothes or other forms of life necessities, volunteered, or offered help to strangers.

Financial behavior: Respondents were asked to indicate to what frequency they run an outstanding balance and pay financial charges, pay the minimum payment only, get charged a late fee, get charged an overcharge fee, use cards for cash advances, have accounts closed down by the bank or credit card companies, or pay credit cards in full.

In addition to the demographic variables, the measures described below were also assessed to serve as criterion variables:

Big Five Inventory-2 Extra Short Form (Soto & John, 2017). This is a 15-item measure of the Big Five personality domains, with three items per domain. Its scales are Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness. Higher scores indicate a higher standing on the trait as described by the scale name.

Dark Triad (Johason & Webster, 2010. This 12-item scale measures the Dark Triad, namely Narcissism, Psychopathy, and Machiavellianism.

Moral Disengagement (Moore, Detert, Klebe, Trevino, Baker, & Mayer, 2012). It has been argued that people can behave unethically when their moral self-regulatory processes are deactivated. The Moral Disengagement scales includes eight items that assess processes that may

be used for this deactivation process. An example item is “People who get mistreated have usually done something to bring it on themselves.”

Subjective Well-being (SWB; Diener, Emmons, Larsen, & Griffin, 1985). This five-item scale measures people’s feelings about their lives. A sample item is “In most ways my life is close to ideal.”

Organizational Citizenship Behavior (OCB; Spector, Bauer, & Fox, 2010). This ten-item scale measures various ways that employees can contribute more to their organization than is in their job description. An example item is “Helped new employees get oriented to the job.”

Counterproductive Work Behavior (CWB; Spector, Bauer, & Fox, 2010). In contrast to OCB, this ten-item scale assess behaviors that are destructive to one’s organization. An example item is “Complained about insignificant things at work.”

### 3.3 Samples

The two forms (single statement and 2AFC) were administered to samples of 500 MTurk workers using honest and fake good instructions. The single statement version and the 2AFC version were counterbalanced in order. Although 431 respondents in the honest condition and 418 respondents in the fake good condition correctly answered at least six of the seven quality control items, the cleaned data sets were nonetheless anomalous. For example, after reverse scoring the Likert items, substantial negative correlations between positive and negative items were observed for the TAPAS Bright Side scales; these scales have been used in the past without this problem. OCB and CWB, two conceptually independent if not negatively related constructs, had a large positive correlation.

To explore this aberrant responding, the Mahalanobis distance was used for outlier detection. It is defined as

$$D_i^2 = (x_i - \overline{x})' S^{-1} (x_i - \overline{x})$$

where $x_i$ is a vector of item responses by the *i*th person, $\overline{x}$ is a vector of item means, and *S* is the variance-covarance matrix of the items. The vector $\overline{x}$ and matrix $S$ were based on an MTurk sample that showed normal results and was collected in May 2019; variables included the 15 BFI-2-S ote,s the 10 CWB items and the 10 OCB items. Large values of the Mahalanobis distance indicate outliers. We began by deleting cases with very large distances ( $D_i^2 \geq 120$) and recomputing the item-item correlation matrices for Bright Side scales and the OCB-CWB correlation. Even after deleting the most extreme cases, these correlations remained anomalous. We iteratively repeated this analysis for smaller ( $D_i^2 \geq 110$) and smaller ( $D_i^2 \geq 100$) until the correlations became plausible. At this point we were deleting cases with $D_i^2 \geq 50$, and 109 and 102 cases remained in the honest and fake good conditions.

Due to the large proportion of “bad data”, we began a search for other on-line crowdsourcing platforms. We found that Prolific was reputed to yield high quality data, albeit at a somewhat higher cost than MTurk. We did an initial test with a sample of 100 and found the Prolific data to

be consistent with what we have found in the past (e.g., after reverse scoring negative items, all items on each Bright Side scale were positively correlated).

Given the positive findings for Prolific, we used this platform to collect 331 additional cases in the honest condition and 464 cases in the fake good condition. After deleting Prolific cases that failed the quality control checks, and combining with the relative few good cases from MTurk, we had 504 honest and 478 fake good cases for analysis.

### 3.4 Results

To interpret results, note that:

- Higher scores on the bright traits mean "good," (i.e., a high score on Selflessness means the respondent reported a great deal of selfless behavior)
- Higher scores on the dark traits mean "bad," (i.e., a high score on Machiavellianism means the respondent reported a great deal of scheming and plotting to the detriment of others).

Consequently, faking good should inflate scores on the Bright side dimensions and deflate scores on the Dark side dimensions.

#### 3.4.1. Descriptive Statistics

Table 1 presents the descriptive statistics for the SS and 2AFC format scales. Scores were computed as the mean response and, for the SS scales, reliability was assessed via coefficient alpha. Reliabilities in the Honest condition were moderate, ranging from .63 for Machiavellianism to .76 for Narcissism for the Dark Side traits. For Bright Side traits, reliabilities ranged from .65 for Selflessness to .71 for Achievement and Even-Tempered. Interestingly, the reliabilities for the SS scales were much higher in the Faking condition, ranging from .87 for Machiavellianism to .93 for Narcissism. One interpretation of this difference is that in the Honest condition, respondents were carefully responding to the items, which ask about the trait from various perspectives. Consequently, using terminology from factor analysis, each item contained common variance (i.e., the trait assessed by all the items on a scale), specific variance (i.e., reliable variance specific to each item), and random error variance. The moderate reliabilities in the Honest condition suggest that there was substantial specific variance for each item, which resulted from the respondents' careful consideration of what the item asked. On the other hand, the very high reliabilities in the Faking condition suggest that respondents ignored the idiosyncratic aspect of each item and responded to all items based on their conception of what an ideal employee would look like.

Note that the means of the Bright Side traits in the Faking condition were considerably elevated from the Honest condition. The t-statistics are very large and significant at any conceivable p-value. The effect sizes range from .63 for Achievement and Selflessness to .74 for Virtue. These are somewhat larger than given in Viswesvaran and Ones's (1999) meta-analysis of instructed faking studies. The means for all four Dark Side traits differed significantly across the Faking and Honest conditions, albeit with much smaller t-statistics. Three of the four were significantly lower, as expected, but Narcissism was significantly higher. Apparently, the respondents

believed that the self-importance, attention seeking, and lack of humility behaviors of narcissists are desirable.

Coefficient alpha cannot be computed for the 2AFC items, so IRT marginal reliability was used as the estimate of reliability,

$$\text{Marginal reliability} = 1 - \frac{\overline{\sigma}_e^2}{\sigma_{\hat{\theta}}^2} \ ,$$

where $\overline{\sigma}_e^2$ is the average squared standard error of the latent trait estimate $\hat{\theta}$ and $\sigma_{\hat{\theta}}^2$ is the variance of the $\hat{\theta}$ values.

The reliability estimates for the 2AFC scales in Table 1 are low, ranging from .36 for Sadism to .57 for Narcissism in the Honest condition and .44 for Achievement to .64 for Narcissism in the Faking condition. These reliabilities are approximately in the range found for the multidimensional 2AFC AF TAPAS, which measures 15 Bright Side traits (e.g., 2AFC Selflessness scale was found to have an IRT reliability of .33 and Achievement was .61). These reliabilities suggest that respondents find that, even in the Faking condition, the 2AFC are difficult to answer and they are not able to answer in a highly consistently manner.

For the 2AFC format, the means of the Bright Side traits were higher in the Faking condition , albeit not as much higher as for the SS scales. The effect sizes ranged from .36 for Selflessness to .67 for Virtue. For the Dark Side traits, Machiavellianism was significantly, but only slightly, lower (effect size -0.14), Psychopathy did not differ across the two conditions, and Narcissism and Sadism were significantly higher in the Faking condition.

**Table 1. Descriptive Statistics for the Dark Side AF TAPAS**

| | Honest (N = 504) | | | Faking (N = 478) | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | M | SD | Reliability | M | SD | Reliability | t-statistic | p-value | Cohen's d |
| SS Achievement | 3.53 | 0.69 | 0.71 | 3.98 | 0.74 | 0.88 | 9.86 | 0.00 | 0.63 |
| SS Even Tempered | 3.43 | 0.82 | 0.71 | 3.95 | 0.77 | 0.88 | 10.21 | 0.00 | 0.65 |
| SS Selflessness | 3.42 | 0.60 | 0.65 | 3.83 | 0.71 | 0.88 | 9.79 | 0.00 | 0.63 |
| SS Virtue | 3.24 | 0.66 | 0.66 | 3.76 | 0.76 | 0.88 | 11.52 | 0.00 | 0.74 |
| SS Mach | 2.17 | 0.64 | 0.63 | 1.91 | 0.70 | 0.87 | -6.23 | 0.00 | -0.40 |
| SS Narcissism | 2.68 | 0.60 | 0.76 | 2.79 | 0.58 | 0.93 | 2.93 | 0.00 | 0.19 |
| SS Psychopathy | 2.18 | 0.44 | 0.68 | 2.10 | 0.51 | 0.89 | -2.49 | 0.01 | -0.16 |
| SS Sadism | 2.45 | 0.68 | 0.68 | 2.28 | 0.71 | 0.89 | -3.76 | 0.00 | -0.24 |
| FC Achievement | 0.29 | 0.84 | 0.42 | 0.73 | 0.81 | 0.44 | 8.33 | 0.00 | 0.53 |
| FC Even Tempered | -0.63 | 0.80 | 0.46 | -0.27 | 0.59 | 0.49 | 8.04 | 0.00 | 0.51 |
| FC Selflessness | -0.44 | 0.64 | 0.42 | -0.22 | 0.63 | 0.46 | 5.56 | 0.00 | 0.36 |
| FC Virtue | -1.00 | 0.57 | 0.48 | -0.62 | 0.58 | 0.47 | 10.42 | 0.00 | 0.67 |
| FC Mach | -1.13 | 0.73 | 0.42 | -1.23 | 0.68 | 0.46 | -2.24 | 0.03 | -0.14 |
| FC Narcissism | -0.57 | 1.12 | 0.57 | -0.09 | 1.42 | 0.64 | 5.82 | 0.00 | 0.37 |
| FC Psychopathy | -1.30 | 0.54 | 0.43 | -1.31 | 0.54 | 0.47 | -0.26 | 0.80 | -0.02 |
| FC Sadism | -0.02 | 0.89 | 0.36 | 0.25 | 0.84 | 0.54 | 4.82 | 0.00 | 0.31 |

*Note*: SS = single statement; FC = two-alternative forced choice.

### 3.4.2. Cross-Format Correlations

Table 2 presents the convergent and discriminant validity correlations for responses obtained in the Honest condition. For the Bright Side scales, these range from the fairly large .61 for Even Tempered scales to .47 for Selflessness. For the Dark Side traits, the convergent correlations range from .21 for Narcissism to .44 for Psychopathy. Clearly, it is more difficult to assess the Dark Side traits.

**Table 2. Cross-Format Correlations Obtained in the Honest Condition**

| | SS Achievement | SS Even Tempered | SS Selflessness | SS Virtue | SS Mach | SS Narcissism | SS Psychopathy | SS Sadism |
|---|---|---|---|---|---|---|---|---|
| FC Achievement | **0.56** | 0.14 | 0.05 | 0.08 | -0.01 | 0.31 | 0.05 | 0.03 |
| FC Even Tempered | 0.08 | **0.61** | -0.09 | 0.03 | 0.03 | 0.16 | 0.14 | 0.05 |
| FC Selflessness | 0.14 | 0.08 | **0.47** | 0.17 | -0.11 | 0.11 | -0.12 | -0.13 |
| FC Virtue | 0.09 | 0.05 | 0.13 | **0.50** | -0.15 | 0.14 | 0.03 | -0.01 |
| FC Mach | -0.06 | -0.15 | -0.13 | -0.19 | **0.35** | 0.08 | 0.09 | 0.18 |
| FC Narcissism | 0.25 | 0.28 | 0.24 | 0.29 | -0.16 | **0.21** | -0.10 | -0.06 |
| FC Psychopathy | -0.07 | -0.10 | -0.26 | -0.10 | 0.18 | 0.09 | **0.44** | 0.35 |
| FC Sadism | 0.06 | 0.05 | -0.07 | 0.02 | 0.01 | 0.08 | 0.13 | **0.40** |

*Note:* SS = single statement; FC = two-alternative forced-choice; convergent cross-format correlations are in bold.

As expected, Table 2 shows that Machiavellianism is negatively correlated with Selflessness and Virtue: SS Machiavellianism correlates -.11 with 2AFC Selflessness and -.15 with 2AFC Virtue and 2AFC Machiavellianism correlates -.13 with SS Selflessness and -.19 with SS Virtue.

For comparison, Table 3 presents the convergent and discriminant validity correlations obtained under the Faking condition. The observed correlations, as one might expect, are generally smaller than those observed in the Honest condition, especially for the Bright Side traits.

**Table 3. Cross-Format Correlations Obtained in the Faking Condition**

| | SS Achievement | SS Even Tempered | SS Selflessness | SS Virtue | SS Mach | SS Narcissism | SS Psychopathy | SS Sadism |
|---|---|---|---|---|---|---|---|---|
| FC Achievement | **0.39** | 0.23 | 0.21 | 0.23 | -0.18 | 0.11 | -0.11 | -0.24 |
| FC Even Tempered | 0.10 | **0.35** | 0.13 | 0.10 | -0.08 | 0.00 | -0.09 | -0.09 |
| FC Selflessness | 0.23 | 0.20 | **0.37** | 0.27 | -0.18 | -0.01 | -0.20 | -0.21 |
| FC Virtue | 0.27 | 0.18 | 0.27 | **0.45** | -0.28 | 0.08 | -0.14 | -0.22 |
| FC Mach | -0.23 | -0.30 | -0.27 | -0.31 | **0.40** | 0.09 | 0.29 | 0.28 |
| FC Narcissism | 0.37 | 0.33 | 0.35 | 0.41 | -0.34 | **0.14** | -0.24 | -0.26 |
| FC Psychopathy | -0.21 | -0.25 | -0.26 | -0.19 | 0.25 | 0.09 | **0.37** | 0.32 |
| FC Sadism | 0.06 | 0.04 | -0.05 | 0.04 | 0.04 | 0.11 | 0.08 | **0.26** |

### 3.4.3. Correlations with Criterion Variables

In the final set of analyses, the scales were correlated with the criterion variables. These correlations appear in Tables 4 for the Honest condition and Table 5 for the Faking condition.

Table 4 shows convergent validity for the Bright Side scales and the three Dark Triad scales. For example, achievement and virtue are facets of the Big Five achievement domain; SS Achievement and SS Virtue correlate .52 and .39 with the BFI-2 Conscientiousness scale. For the 2AFC format, Achievement and Virtue correlate .32 and .24 with BFI-2 Conscientiousness. Even tempered is a facet of Neuroticism, and SS and 2AFC Even Tempered correlate -0.62 and -0.51 with BFI-2 Neuroticism. Selflessness is a facet of the Big Five agreeableness domain; SS and 2AFC Selflessness correlate .49 and .24 with BFI-2 Agreeableness.

Turning now to the Dark Side, the SS and 2AFC Machiavellianism scales correlated 0.66 and 0.31 with the Dark Triad Machiavellianism scale. The SS and 2AFC Psychopathy scales correlated .60 and .36 with the Dark Triad Psychopathy scale. The SS Narcissism scale correlated .66 with its Dark Triad counterpart, but the 2AFC Narcissism scale had only a .12 correlation.

The Bright Side SS and 2AFC scales had good correlations with the positive criterion variables SWB and OCB. For example, SS and 2AFC Achievement had .31 and .24 correlations with SWB and .35 fand .26 correlations with OCB. SS and 2AFC Virtue similarly correlated with SWB (.28 and .25) and OCB (.26 and .16). On the other hand, SS and 2AFC Achievement and

**Table 4. Scale Correlations for the Honest Condition**

| | Achievement | | Even Tempered | | Selflessness | | Virtue | | Machiavellianism | | Narcissism | | Psychopathy | | Sadism | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC |
| Education | 0.22 | 0.17 | 0.08 | 0.08 | 0.12 | 0.15 | 0.16 | 0.17 | -0.03 | 0.03 | 0.16 | 0.14 | 0.03 | 0.04 | 0.02 | 0.06 |
| Income | 0.20 | 0.17 | 0.07 | 0.06 | 0.12 | 0.14 | 0.03 | -0.05 | -0.07 | -0.09 | 0.08 | 0.05 | -0.11 | 0.00 | -0.06 | -0.03 |
| Employment | -0.23 | -0.18 | -0.11 | -0.13 | -0.08 | -0.10 | -0.19 | -0.20 | 0.02 | -0.02 | -0.20 | -0.12 | -0.06 | -0.13 | -0.12 | -0.14 |
| Age | 0.11 | 0.09 | 0.07 | 0.04 | 0.10 | 0.10 | 0.22 | 0.16 | -0.07 | 0.04 | 0.03 | 0.09 | -0.03 | -0.06 | 0.01 | 0.02 |
| Status | 0.21 | 0.22 | 0.09 | 0.20 | 0.05 | 0.21 | 0.13 | 0.24 | 0.07 | 0.02 | 0.31 | 0.18 | 0.12 | 0.14 | 0.14 | 0.11 |
| SocMedia | 0.19 | 0.08 | 0.02 | -0.01 | 0.16 | 0.13 | 0.14 | 0.18 | -0.04 | 0.08 | 0.22 | 0.15 | -0.05 | 0.08 | 0.00 | 0.09 |
| Charity | 0.25 | 0.22 | 0.04 | 0.08 | 0.33 | 0.32 | 0.23 | 0.29 | -0.03 | 0.01 | 0.26 | 0.18 | -0.06 | 0.06 | 0.02 | 0.06 |
| FinaBehav | -0.03 | -0.07 | -0.03 | -0.04 | -0.11 | -0.11 | -0.12 | -0.24 | -0.04 | -0.02 | -0.10 | -0.12 | -0.08 | -0.03 | -0.09 | -0.08 |
| Extraversion | 0.43 | 0.33 | 0.20 | 0.16 | 0.21 | 0.21 | 0.23 | 0.22 | -0.10 | -0.03 | 0.53 | 0.27 | -0.05 | 0.09 | 0.09 | 0.21 |
| Agreeable | 0.28 | 0.09 | 0.43 | 0.12 | **0.49** | **0.24** | 0.39 | 0.23 | -0.44 | -0.12 | 0.14 | 0.25 | -0.48 | -0.26 | -0.35 | -0.07 |
| Conscientious | **0.52** | **0.32** | 0.35 | 0.17 | 0.22 | 0.08 | **0.39** | **0.24** | -0.28 | -0.02 | 0.24 | 0.21 | -0.15 | 0.03 | -0.06 | 0.13 |
| Neuroticism | -0.36 | -0.23 | **-0.62** | **-0.51** | -0.13 | -0.07 | -0.31 | -0.20 | 0.23 | 0.08 | -0.26 | -0.27 | -0.08 | -0.12 | -0.06 | -0.20 |
| Openness | 0.30 | 0.08 | 0.14 | -0.05 | 0.28 | 0.08 | 0.16 | -0.02 | -0.20 | 0.00 | 0.10 | 0.16 | -0.27 | -0.12 | -0.15 | 0.00 |
| DTNAR | 0.18 | 0.23 | 0.03 | 0.05 | -0.02 | 0.08 | -0.01 | 0.16 | 0.18 | 0.02 | **0.66** | **0.12** | -0.06 | 0.06 | 0.11 | 0.06 |
| DTPSY | -0.21 | 0.03 | -0.26 | 0.06 | -0.46 | -0.09 | -0.35 | -0.02 | 0.49 | 0.19 | 0.11 | -0.10 | **0.60** | **0.36** | 0.52 | 0.12 |
| DTMACH | -0.15 | 0.05 | -0.25 | 0.00 | -0.32 | -0.12 | -0.41 | -0.13 | **0.66** | **0.31** | 0.17 | -0.10 | 0.27 | 0.19 | 0.34 | 0.12 |
| MoralDis | -0.09 | 0.04 | -0.16 | 0.10 | -0.29 | 0.00 | -0.26 | 0.09 | 0.44 | 0.19 | 0.30 | -0.01 | 0.38 | 0.23 | 0.41 | 0.17 |
| SWB | 0.31 | 0.24 | 0.29 | 0.21 | 0.22 | 0.18 | 0.28 | 0.25 | -0.19 | -0.08 | 0.29 | 0.28 | -0.11 | 0.06 | -0.09 | 0.10 |
| OCB | 0.35 | 0.26 | 0.13 | 0.09 | 0.26 | 0.22 | 0.26 | 0.16 | -0.14 | 0.00 | 0.17 | 0.16 | -0.03 | 0.03 | 0.01 | -0.01 |
| CWB | -0.15 | -0.01 | -0.13 | 0.01 | -0.15 | 0.05 | -0.11 | 0.02 | 0.32 | 0.09 | 0.12 | 0.03 | 0.22 | 0.18 | 0.24 | 0.10 |

*Note:* observed correlations with absolute values greater than 0.09 are significant at p = .05 (one-tailed) and greater than 0.12 are significant at p = .01 (one-tailed). FinaBehav = Financial Behavior, Extraversion = BFI-2 Extraversion, Agreeableness = BFI-2 Agreeableness, Conscientious = BFI-2 Conscientiousness, Neuroticism = BFI-2 Neuroticism, Openness = BFI-2 Openness, DTNAR = Dark Triad Narcissism, DTPSY = Dark Triad Psychopathy, DTMACH = Dark Triad Machiavellianism, MoralDis = Moral Disengagement, SWB = Subjective Well-Being, OCB = Organizational Citizenship Behavior, CWB = Counterproductive Work Behavior.

Even-Tempered had small correlations with the negative Moral Disengagment and CWB. But SS Selflessness and SS Virtue had -.29 and -.26 correlations with Moral Disengagment.

The Dark Tetrad had much stronger correlations with the negative criteria but mostly small correlations with the positive criteria. For example, SS Machiavellianism, SS Narcissism, SS Psychopathy, and SS Sadism correlated .44, .30, .38 and .41 with Moral Disengagement and SS Machiavellianism had a .32 correlation with CWB.

Table 5 shows the scale correlation for the Faking condition. The general pattern of convergent validities is similar, but the magnitude of the correlations is substantially smaller. For example the SS Dark Triad convergent validities were in the .60s in Table 4, but in the .40s in Table 5. The Bright Side convergent validities are very much smaller in the Faking condition. Interestingly, for the Faking condition the four Bright Side SS scales had correlations that ranged from -.39 to -.44 with Moral Disengagement and SS Machiavellianism had a .56 correlation with this criterion variable.

**Perhaps the most salient feature of these tables is that scale scores computed from responses in the Faking condition have little predictive utility for the criterion variables.** None of the adjusted $R^2$ values was larger than .17, and many were less than .10. In contrast, scale scores from the Honest condition had adjusted $R^2$ values usually in the .2 to .4 range, which indicates fairly good prediction.

## Table 5. Scale Correlations for the Fake Good Condition

| | Achievement | | Even Tempered | | Selflessness | | Virtue | | Machiavellianism | | Narcissism | | Psychopathy | | Sadism | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC |
| Education | -0.13 | 0.02 | -0.16 | -0.08 | -0.08 | 0.02 | -0.09 | 0.05 | 0.14 | 0.14 | 0.15 | -0.07 | 0.07 | 0.07 | 0.18 | 0.01 |
| Income | 0.10 | -0.02 | 0.07 | -0.02 | 0.13 | 0.02 | 0.10 | -0.01 | -0.07 | 0.02 | 0.01 | 0.05 | -0.16 | -0.05 | -0.04 | 0.09 |
| Employment | 0.08 | 0.03 | 0.06 | 0.04 | 0.05 | 0.02 | 0.05 | -0.05 | -0.10 | -0.19 | -0.14 | 0.01 | -0.04 | -0.09 | -0.16 | -0.10 |
| Age | -0.05 | -0.13 | -0.04 | -0.09 | -0.02 | -0.08 | -0.03 | 0.03 | 0.05 | 0.09 | 0.03 | 0.00 | 0.07 | 0.07 | 0.12 | 0.11 |
| Status | -0.16 | -0.06 | -0.21 | -0.06 | -0.17 | -0.06 | -0.12 | 0.05 | 0.22 | 0.21 | 0.20 | -0.03 | 0.16 | 0.12 | 0.24 | 0.10 |
| SocMedia | 0.00 | -0.02 | -0.08 | -0.07 | 0.00 | 0.05 | -0.01 | 0.09 | 0.08 | 0.17 | 0.19 | -0.02 | 0.00 | 0.15 | 0.07 | 0.09 |
| Charity | -0.11 | -0.01 | -0.16 | -0.07 | -0.01 | 0.03 | -0.08 | 0.13 | 0.17 | 0.17 | 0.26 | -0.01 | 0.07 | 0.08 | 0.17 | 0.12 |
| FinaBehav | 0.27 | 0.12 | 0.26 | 0.03 | 0.22 | 0.06 | 0.23 | -0.09 | -0.32 | -0.09 | -0.16 | 0.10 | -0.19 | -0.08 | -0.31 | -0.06 |
| Extraversion | -0.07 | -0.07 | -0.15 | -0.02 | -0.11 | -0.03 | -0.10 | 0.06 | 0.14 | 0.10 | 0.36 | 0.01 | 0.08 | 0.13 | 0.28 | 0.06 |
| Agreeable | 0.10 | -0.07 | 0.16 | 0.01 | **0.22** | **0.03** | 0.22 | 0.04 | -0.27 | -0.07 | 0.07 | 0.10 | -0.27 | -0.17 | -0.16 | -0.03 |
| Conscientious | **0.16** | **0.02** | 0.11 | -0.01 | 0.13 | -0.07 | **0.12** | **0.09** | -0.12 | 0.10 | 0.17 | 0.14 | -0.16 | 0.01 | -0.02 | 0.06 |
| Neuroticism | -0.01 | 0.07 | **-0.13** | **-0.12** | 0.01 | 0.10 | 0.00 | 0.00 | 0.01 | -0.05 | -0.20 | -0.11 | 0.00 | -0.01 | -0.11 | -0.09 |
| Openness | 0.26 | 0.07 | 0.17 | 0.00 | 0.19 | 0.08 | 0.15 | 0.05 | -0.18 | -0.03 | 0.03 | 0.08 | -0.18 | -0.03 | -0.13 | 0.02 |
| DTNAR | -0.10 | 0.07 | -0.18 | -0.07 | -0.13 | -0.03 | -0.15 | 0.09 | 0.23 | 0.12 | **0.49** | **-0.03** | 0.07 | 0.04 | 0.20 | 0.08 |
| DTPSY | -0.27 | 0.02 | -0.36 | -0.02 | -0.38 | -0.01 | -0.33 | 0.00 | 0.43 | 0.16 | 0.09 | -0.16 | **0.43** | **0.20** | 0.33 | 0.11 |
| DTMACH | -0.17 | 0.07 | -0.25 | -0.03 | -0.23 | -0.03 | -0.29 | -0.08 | **0.43** | **0.19** | 0.14 | -0.07 | 0.26 | 0.15 | 0.25 | 0.09 |
| MoralDis | -0.39 | -0.03 | -0.44 | -0.01 | -0.43 | -0.02 | -0.43 | -0.02 | 0.56 | 0.22 | 0.24 | -0.19 | 0.39 | 0.16 | 0.43 | 0.08 |
| SWB | -0.08 | -0.01 | -0.09 | 0.01 | -0.03 | 0.02 | 0.01 | 0.10 | 0.08 | 0.10 | 0.21 | -0.02 | 0.01 | 0.04 | 0.10 | -0.03 |
| OCB | -0.03 | 0.02 | -0.03 | -0.06 | -0.01 | -0.01 | 0.05 | 0.09 | 0.01 | -0.02 | 0.09 | 0.01 | 0.03 | 0.02 | 0.08 | 0.01 |
| CWB | -0.21 | -0.04 | -0.19 | -0.01 | -0.19 | 0.00 | -0.16 | -0.01 | 0.28 | 0.09 | 0.16 | -0.09 | 0.22 | 0.13 | 0.21 | 0.06 |

### 3.4.4. Correlations with Criterion Variables Corrected for Unreliability in the Criteria

The Big Five Inventory-2 Extra Short Form, the Dark Triad, Moral Disengagement, SWB, OCB, and CWB are all measured with error. As these measures are serving as criterion variables, it is interesting to examine the extent to which the new scales predict their true scores. To this end, we computed the SS and 2AFC scales' correlations with the criterion measures correcting for unreliability in the criteria. Table 6 shows the corrected correlations for the Honest condition and Table 7 shows the corrected correlations for the Faking condition. These tables show the same general trends as Tables 4 and 5, albeit with somewhat larger values.

**Table 6. Scale Correlations for the Honest Condition after Correcting for Unreliability in the Criteria**

| | Achievement | | Even-Tempered | | Selflessness | | Virtue | | Machiavellianism | | Narcissism | | Psychopathy | | Sadism | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC |
| Extraversion | 0.52 | 0.40 | 0.24 | 0.20 | 0.25 | 0.25 | 0.28 | 0.26 | -0.12 | -0.04 | 0.64 | 0.33 | -0.06 | 0.11 | 0.10 | 0.25 |
| Agreeable | 0.34 | 0.11 | 0.53 | 0.15 | 0.59 | 0.29 | 0.47 | 0.28 | -0.53 | -0.14 | 0.17 | 0.30 | -0.58 | -0.32 | -0.42 | -0.09 |
| Conscientious | 0.63 | 0.39 | 0.43 | 0.21 | 0.27 | 0.10 | 0.48 | 0.30 | -0.35 | -0.02 | 0.30 | 0.26 | -0.19 | 0.04 | -0.07 | 0.17 |
| Neuroticism | -0.44 | -0.28 | -0.76 | -0.62 | -0.16 | -0.08 | -0.38 | -0.25 | 0.28 | 0.10 | -0.32 | -0.33 | -0.10 | -0.14 | -0.08 | -0.24 |
| Openness | 0.35 | 0.09 | 0.17 | -0.05 | 0.33 | 0.10 | 0.19 | -0.02 | -0.24 | 0.00 | 0.12 | 0.19 | -0.32 | -0.14 | -0.17 | 0.00 |
| DTNAR | 0.21 | 0.27 | 0.04 | 0.06 | -0.02 | 0.10 | -0.01 | 0.19 | 0.21 | 0.03 | 0.78 | 0.14 | -0.08 | 0.07 | 0.13 | 0.07 |
| DTPSY | -0.25 | 0.04 | -0.31 | 0.07 | -0.55 | -0.10 | -0.42 | -0.02 | 0.58 | 0.23 | 0.13 | -0.12 | 0.72 | 0.43 | 0.62 | 0.15 |
| DTMACH | -0.18 | 0.06 | -0.29 | 0.00 | -0.38 | -0.14 | -0.49 | -0.16 | 0.79 | 0.37 | 0.21 | -0.12 | 0.32 | 0.22 | 0.40 | 0.14 |
| MoralDis | -0.11 | 0.05 | -0.19 | 0.12 | -0.33 | 0.00 | -0.30 | 0.10 | 0.52 | 0.23 | 0.35 | -0.01 | 0.44 | 0.27 | 0.48 | 0.20 |
| SWB | 0.38 | 0.29 | 0.35 | 0.25 | 0.27 | 0.22 | 0.34 | 0.30 | -0.24 | -0.09 | 0.35 | 0.34 | -0.13 | 0.07 | -0.11 | 0.12 |
| OCB | 0.41 | 0.31 | 0.15 | 0.11 | 0.31 | 0.26 | 0.31 | 0.19 | -0.17 | 0.00 | 0.20 | 0.19 | -0.04 | 0.04 | 0.02 | -0.01 |
| CWB | -0.17 | -0.01 | -0.16 | 0.01 | -0.17 | 0.06 | -0.13 | 0.03 | 0.38 | 0.11 | 0.14 | 0.03 | 0.27 | 0.21 | 0.28 | 0.12 |

**Table 7. Scale Correlations for the Fake Good Condition after Correcting for Unreliability in the Criteria**

| | Achievement | | Even Tempered | | Selflessness | | Virtue | | Machiavellianism | | Narcissism | | Psychopathy | | Sadism | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC | SS | 2AFC |
| Extraversion | -0.08 | -0.08 | -0.18 | -0.02 | -0.13 | -0.04 | -0.12 | 0.06 | 0.16 | 0.11 | 0.43 | 0.01 | 0.10 | 0.16 | 0.33 | 0.07 |
| Agreeable | 0.12 | -0.08 | 0.19 | 0.02 | 0.26 | 0.04 | 0.26 | 0.05 | -0.32 | -0.08 | 0.08 | 0.11 | -0.32 | -0.20 | -0.18 | -0.04 |
| Conscientious | 0.19 | 0.03 | 0.13 | -0.01 | 0.16 | -0.09 | 0.14 | 0.10 | -0.14 | 0.12 | 0.20 | 0.16 | -0.19 | 0.01 | -0.03 | 0.07 |
| Neuroticism | -0.01 | 0.09 | -0.16 | -0.14 | 0.01 | 0.12 | 0.00 | 0.00 | 0.01 | -0.06 | -0.23 | -0.13 | 0.00 | -0.02 | -0.14 | -0.11 |
| Openness | 0.30 | 0.08 | 0.19 | 0.00 | 0.22 | 0.09 | 0.18 | 0.06 | -0.21 | -0.03 | 0.04 | 0.09 | -0.21 | -0.04 | -0.16 | 0.02 |
| DTNAR | -0.12 | 0.08 | -0.21 | -0.08 | -0.15 | -0.03 | -0.17 | 0.10 | 0.27 | 0.13 | 0.56 | -0.04 | 0.08 | 0.04 | 0.23 | 0.09 |
| DTPSY | -0.32 | 0.03 | -0.42 | -0.02 | -0.44 | -0.02 | -0.39 | 0.00 | 0.50 | 0.19 | 0.10 | -0.19 | 0.50 | 0.24 | 0.39 | 0.12 |
| DTMACH | -0.20 | 0.08 | -0.30 | -0.04 | -0.26 | -0.03 | -0.33 | -0.09 | 0.51 | 0.22 | 0.17 | -0.08 | 0.30 | 0.17 | 0.30 | 0.11 |
| MoralDis | -0.45 | -0.03 | -0.51 | -0.01 | -0.49 | -0.02 | -0.49 | -0.03 | 0.64 | 0.25 | 0.27 | -0.22 | 0.44 | 0.19 | 0.50 | 0.09 |
| SWB | -0.09 | -0.01 | -0.11 | 0.01 | -0.03 | 0.02 | 0.01 | 0.11 | 0.10 | 0.12 | 0.25 | -0.02 | 0.01 | 0.04 | 0.12 | -0.04 |
| OCB | -0.03 | 0.02 | -0.03 | -0.07 | -0.02 | -0.01 | 0.05 | 0.10 | 0.01 | -0.02 | 0.10 | 0.02 | 0.03 | 0.03 | 0.09 | 0.01 |
| CWB | -0.24 | -0.04 | -0.22 | -0.02 | -0.22 | 0.00 | -0.18 | -0.01 | 0.32 | 0.10 | 0.18 | -0.10 | 0.25 | 0.15 | 0.24 | 0.07 |

### 3.4.5. Adjusted Squared Multiple Correlations

In the final set of analyses, adjusted (to correct for capitalization on chance) $R^2$ values were computed for each scale format in each administrative condition. The entries in Table 8 result from predicting each variable in the first column from the eight scales included in each format-administrative condition combination. It is clear that instructions to fake good reduced the validity of the personality scales for predicting the demographic and criterion variables. For the Honest condition, the SS and 2AFC formats have generally comparable predictive power for predicting the demographic variables (e.g., $R^2$s of .23 and .20 for Charity), but the SS has far more predictive power for predicting the Likert formatted self-ratings (e.g., the BFI-2 Extraversion, Agreeable, Conscientious, Neuroticism, Openness). These latter $R^2$s may be inflated due to common method variance. Interestingly, the Dark Triad scales (DTNAR for narcissism, DTPSY for psychopathy, and DTMACH for Machiavellianism) are better predicted by the SS scales in the Fake Good condition than the 2AFC scales in either condition.

**Table 8. Adjusted $R^2$ Values**

| | Honest | | Fake Good | |
|---|---|---|---|---|
| | SS | 2AFC | SS | 2AFC |
| Education | 0.08 | 0.06 | 0.06 | 0.02 |
| Income | 0.04 | 0.04 | 0.02 | 0.00 |
| Employment | 0.12 | 0.10 | 0.03 | 0.03 |
| Age | 0.05 | 0.03 | 0.00 | 0.02 |
| Status | 0.16 | 0.17 | 0.09 | 0.04 |
| SocMedia | 0.08 | 0.07 | 0.04 | 0.05 |
| Charity | 0.23 | 0.20 | 0.13 | 0.06 |
| FinaBehav | 0.06 | 0.06 | 0.14 | 0.04 |
| Extraversion | 0.41 | 0.25 | 0.18 | 0.01 |
| Agreeable | 0.44 | 0.18 | 0.12 | 0.04 |
| Conscientious | 0.36 | 0.18 | 0.06 | 0.03 |
| Neuroticism | 0.52 | 0.37 | 0.09 | 0.04 |
| Openness | 0.14 | 0.04 | 0.05 | 0.00 |
| DTNAR | 0.46 | 0.08 | 0.28 | 0.04 |
| DTPSY | 0.48 | 0.15 | 0.22 | 0.07 |
| DTMACH | 0.45 | 0.14 | 0.21 | 0.05 |
| MoralDis | 0.32 | 0.10 | 0.33 | 0.08 |
| SWB | 0.20 | 0.18 | 0.06 | 0.01 |
| OCB | 0.18 | 0.11 | 0.01 | 0.00 |
| CWB | 0.13 | 0.03 | 0.09 | 0.01 |

Table 9 shows the Adjusted $R^2$ values after correcting for unreliability in the criteria. It shows the same general pattern as Table 8, but with somewhat larger values.

**Table 9. Adjusted $R^2$ Values After Correcting Criteria for Unreliability**

| | Honest | | Fake Good | |
|---|---|---|---|---|
| | SS | 2AFC | SS | 2AFC |
| Extraversion | 0.59 | 0.36 | 0.27 | 0.02 |
| Agreeable | 0.62 | 0.26 | 0.17 | 0.05 |
| Conscientious | 0.51 | 0.26 | 0.08 | 0.04 |
| Neuroticism | 0.74 | 0.53 | 0.13 | 0.05 |
| Openness | 0.19 | 0.05 | 0.08 | 0.00 |
| DTNAR | 0.65 | 0.11 | 0.39 | 0.05 |
| DTPSY | 0.65 | 0.21 | 0.29 | 0.10 |
| DTMACH | 0.60 | 0.18 | 0.28 | 0.07 |
| MoralDis | 0.44 | 0.14 | 0.45 | 0.11 |
| SWB | 0.29 | 0.26 | 0.09 | 0.02 |
| OCB | 0.26 | 0.16 | 0.02 | 0.00 |
| CWB | 0.17 | 0.04 | 0.12 | 0.02 |

# 4.0 DISCUSSION

## 4.1 Summary

Two new personality inventories were developed. One form is a traditional Likert scale where respondents are instructed to rate each statement on a 5-point scale. The other form is a 2AFC measure where respondents are instructed to choose the statement that is "more like me." The process of scale development began with a literature review to identify conceptualizations of the Dark Tetrad constructs. Following this review, careful definitions of these constructs were developed. Then a large number of statements were drafted. They were reviewed, with duplicates and overly similar statements deleted. Statements deviating too much from the construct definitions were also deleted. The remaining statements were edited to ensure clarity and proper grammar. Finally, about 50 statements per Dark Tetrad constructs were administered to samples of MTurk workers and USAF Basic Recruits.

Analysis of the resulting data proved problematic. In the past we have often had to run GGUM2004 two or three times and delete a few problematic statements in order to obtain convergence to a proper solution for a particular trait. With the Dark Tetrad constructs, we made dozens of runs. We ultimately tried four different estimation programs, including a new Bayesian method. By combining the MTurk and USAF Basic Recruit data and deleting a large number of statements, we were finally able to secure sensible parameter estimates for the statements.

After estimating parameters, two scale versions were assembled. Both included the Dark Tetrad constructs as well as four Bright Side traits (achievement, even tempered, selflessness, and virtue). The single statement (SS) version used the traditional Likert rating scale format with five

rating scale points, The two-alternative forced choice (2AFC) presented respondents with two statements and they must choose the statement that is "more like me."

The SS and 2AFC versions were administered to samples of 500 MTurk workers. In one administrative condition, respondents were instructed to answer honestly and in a second condition they were instructed to fake good. Unfortunately, about 80% of the resulting samples was "bad data". Although most respondents passed the quality control items (i.e., they strongly disagreed with statements like "I can run two miles in two minutes"), they apparently "straight lined" the remaining items by, for example, strongly agreeing with both positively and negatively worded statements for each trait.

We then searched for, and found, an alternative crowd-sourcing platform, Prolific. An initial test with 100 cases yielded positive results, so we collected additional cases to ultimately yield 504 usable cases in the Honest condition and 478 cases in the Fake Good condition.

We then investigated the psychometric properties of the SS and 2AFC scales. The reliabilities of the SS format scales were good, ranging from .63 to .76, with a mean of .69 in the Honest condition. These reliability are in the range of the single statement AF TAPAS; for example the reliability of the SS Selflessness scale was .71 in Chernyshenko et al. (2019). The reliabilities were lower for the forced-choice measures, which are less likely to capitalize on single-subject response consistency error, with a mean of .45 for the Bright Side scales and .45 for the Dark Tetrad scales;. Again, these are similar to the reliabilities found by Chernyshenko et al. for the multidimensional 2AFC AF TAPAS. Reliabilities were implausibly high for the SS scales in the Fake Good condition; for these 8-item scales, coefficient alphas ranged from .87 to .93. IRT reliabilities for the 2AFC scales in the Fake Good condition were very similar to the Honest condition, ranging from .46 to.64.

Effect sizes for the Honest – Fake Good comparisons were large for the Bright Side SS scales, ranging from .63 to .74. They were in the moderate range for the 2AFC format, ranging from .36 to .67. For the Dark Side scales in the 2AFC format, effect sizes were small, ranging from -.14 to .37.

Importantly, effect sizes for the Dark Tetrad scales in the SS format were generally small: -.24 for Sadism, -.16 for Psychopathy, and .19 for Narcissism. Only Machiavellianism approached a moderate effect size, with d = -.40. Apparently, the thoughts, feelings, and behaviors characterizing the Dark Tetrad traits are somewhat confusing to respondents. Whereas it is perfectly clear how to respond to the Bright Side achievement statement "I always get my work done on time" (i.e., Strongly Agree), it is apparently not clear how to respond to the Psychopathy item "I don't get emotional about anything" in the Fake Good condition.

The convergent validity cross-format correlations were found to be reasonably good for the Bright Side scales collected in the Honest condition, but lower for the Dark Tetrad scales. They ranged from a mean correlation of .54 for the Bright Side scales in Table 2 to a mean correlation of .35 for the Dark Tetrad scales. For data collected in the Faking condition, cross-method correlations were much lower, ranging from a mean of .35 to .45 for the Bright Side scales and

from .14 to .40 for the Dark Tetrad scales. After correcting for unreliability in the criteria, these correlations were even higher.

Finally, correlations and squared multiple correlations of the new scales with demographic variables, alternative measures of the traits (the BFI-2 Short Form and the Dark Triad scales), and four criterion variables were examined. Many of the correlations with demographic variables were near zero, but a sensible pattern of larger correlations is apparent (e.g., SS and 2AFC Selflessness correlate .33 and .32 with Charity in the Honest condition). Substantial construct validity was found for the SS scales, as measures of the same or similar construct were generally large (e.g., .60 to .66 for the Dark Triad traits in the Honest condition). The SS scales also had a sensible pattern of correlations with criterion variables, (e.g,. Achievement had a .35 correlation with Organizational Citizenship Behavior (OCB) and Machiavellianism had a .44 correlation with Moral Disengagement). The pattern of correlations was somewhat different in the Fake Good condition, with SS Achievement having a -.03 correlations with OCB but a -.39 correlation with Moral Disengagement. Finally, the 2AFC scales had a similar pattern of correlations as the SS in the Honest condition, although they were consistently lower.

The adjusted $R^2$ values for the SS and 2AFC were comparable for the demographic variables, but the SS values were much higher for the criterion variables using self-report Likert rating scales in the Honest condition. Although Chernyshenko et al. (2019) used a different set of criterion variables, they also observed higher $R^2$ values for the SS scales when the criteria were assessed via the self-report Likert scales. The performance of the 2AFC scales in the Faking condition was disappointing as they showed very little predictive power, again similar to Chernyshenko et al. (2019).

### 4.2 Conclusions and Implications

This project makes it clear that Dark Side traits are psychometrically quite different from Bright Side traits. An inspection of the statements included on a Dark Side SS scale suggests a greater evaluative ambiguity of their thoughts, feelings, and behaviors than observed on a Bright Side scale. One consequence is that faking good is much easier on a SS Bright Side scale: the effect sizes ranged from .63 to .74, with a mean of .66, in Table 1. In contrast, the SS Dark Side effect sizes ranged from -.40 to .19, with a mean of -.15.

This evaluative ambiguity leads to the suggestion that, even in high stakes settings, the SS instrument can be used to assess the Dark Tetrad traits. The four Bright Side scale scores are probably of little use, as Fake Good instructions produce large score increases. But the Fake Good instructions produced little score inflation on the SS Dark Tetrad scales and their construct validity vis-à-vis the Dark Triad scales was moderately strong (correlations ranging from .43 to .49).

## 5.0 REFERENCES


Chabrol, H., Van Leeuwen, N., Rodgers, R., & Séjourné, N. (2009). Contributions of psychopathic, narcissistic, Machiavellian, and sadistic personality traits to juvenile delinquency. *Personality and Individual Differences, 47,* 734–739.

Chernyshenko, O. S., Drasgow, F., Stark, S., & Nye, C.D. (2019). *The development of three static personality research forms and on-line scoring tools for U.S. Air Force*, AFRL-RH-WP-TR-2019-0090. Wright-Patterson AFB, OH: Air Force Research Laboratory, 711 Human Performance Wing, Airman Systems Directorate, Collaborative Interfaces and Teaming Branch.

Costa, P. T., & McCrae, R. R. (2008). The Revised NEO Personality Inventory (NEO-PI-R), In D. H. Saklofske (Ed.), *The SAGE Handbook of Personality Theory and Assessment, Vol. 2: Personality measurement and testing* (pp. 179-198), Thousand Oaks, CA: Sage.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment, 49*, 71–75.

Jacobwitz, S., & Egan, V. (2006). The 'dark triad' and normal personality traits. *Personality and Individual Differences, 40*, 331-339.

Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment, 22*, 420–432

Jones, D. N., & Paulhaus, D. L., (2010). Differentiating the dark triad within the interpersonal circumplex. In L. M. Horowitz & S. N. Strack (Eds.), *Handbook of interpersonal theory and research* (pp. 249-267). NY: Guilford.

Jones, D. N., & Paulhaus, D. L. (2014). Introducing the short dark triad (SD3) a brief measure of dark personality traits. *Assessment, 21*, 28–41.

Kirsch, L. G., & Becker, J. V. (2007). Emotional deficits in psychopathy and sexual sadism: Implications for violent and sadistic behavior. *Clinical Psychology Review, 27,* 904–922.

Kohut, H. (1977). *The restoration of the self.* Chicago, IL: University of Chicago Press.

Mededovic, J., & Petrovic, B. (2015). The dark triad: Structural properties and location in the personality space. *Journal of Individual Differences, 36*, 228-236.

Mokros, A., Osterheider, M., Hucker, S. J., & Nitschke, J. (2011). Psychopathy and sexual sadism. *Law and Human Behavior, 35,* 188–199.

Moore, C., Detert, J. R., Trevino, L. K., Baker, V. L., & Mayer, D. M. (2012). Why employees do bad things: Moral disengagement and unethical organizational behavior. *Personnel Psychology, 65*, 1-48.

Myers, W. C., Burket, R. C., & Husted, D. S. (2006). Sadistic personality disorder and comorbid mental illness in adolescent psychiatric inpatients. *Journal of the American Academy of Psychiatry and the Law Online, 34,* 61–71.

O'Boyle, E. H., Jr., 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.

Paulhaus, D. L, & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality, 36*, 556-563.

Reidy, D. E., Zeichner, A., & Seibert, L. A. (2011). Unprovoked aggression: Effects of psychopathic traits and sadism. *Journal of Personality, 79,* 75–100.

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, 1,* 3-32.

Skeem, J. L., Polaschek, D. L. L., Patrick, P. J., & Lilienfeld, S. O. (2011). Psychopathic personality: Bridging the gap between scientific evidence and public policy. *Psychological Science in the Public Interest, 12*, 95-162.

Soto, C. J., & John, O. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology, 113*, 117-143.

Spector, P. E., Bauer, J. A., & Fox, S. (2010). Measurement artifacts in the assessment of counterproductive work behavior and organizational citizenship behavior: Do we know what we think we know? *Journal of Applied Psychology, 95*, 781–790.

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-unidimensional pairwise-preference model, *Applied Psychological Measurement*, *29, 3,* 184-203

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

White, L. A., & Young, M. C. (1998, August). *Development and validation of the Assessment of Individual Motivation (AIM).* Paper presented at the annual meeting of the American Psychological Association, San Francisco, CA.

## LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS

| | |
|---|---|
| < | Less than |
| 2AFC | Two alternative forced-choice |
| ABLE | Assessment of Background and Life Expereience |
| AIM | Assessment of Individual Motivation |
| BFI-2 Agreeableness | Agreeableness |
| BFI-2 Conscientiousness | Conscientiousness |
| BFI-2 Extraversion | Extraversion |
| BFI-2 Neuroticism | Neuroticism |
| BFI-2 Openness | Openness |
| CTT | Classical Test Theory |
| CWB | Counterproductive Work behavior |
| DTMACH | Dark Triad Machiavellianism |
| DTNAR | Dark Triad Narcissism |
| DTPSY | Dark Triad Psychopathy |
| FinaBehav | Financial Behavior |
| GGUM | Generalized Graded Unfolding Model |
| IRT | Item Response Theory |
| M | Mean |
| MDPP | Multidimensional pairwise preference |
| MoralDis | Moral Disengagement |
| OCB | Organizational Citizenship Behavior |
| R | Multiple Correlation |
| $R^2$ | Variance accounted for (r-squared) |

| | |
|---|---|
| SD | Standard Deviation |
| SocMedia | Social Media |
| SS | Single statement |
| SWB | Subjective well-being |
| TAPAS | Tailored Adaptive Personality Assessment System |
| TSDI | Trait Self-Description Inventory |
| USAF | US Air Force |