

An Examination of the Role of Inverted Dark Tetrad Items on Structural Properties and Construct Validity

Blötner, Christian; Grüning, David J.

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Blötner, C., & Grüning, D. J. (2023). An Examination of the Role of Inverted Dark Tetrad Items on Structural Properties and Construct Validity. *European Journal of Psychological Assessment*. <https://doi.org/10.1027/1015-5759/a000805>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC-ND Licence (Attribution-Non Commercial-NoDerivatives). For more information see:
<https://creativecommons.org/licenses/by-nc-nd/4.0>

An Examination of the Role of Inverted Dark Tetrad Items on Structural Properties and Construct Validity

Christian Blötner¹  and David J. Gruning^{2,3} 

¹Faculty of Psychology, FernUniversität in Hagen, Germany

²Department of Psychology, Heidelberg University, Germany

³Department of Survey Design and Methodology, GESIS – Leibniz Institute for the Social Sciences, Mannheim, Germany

Abstract: The use of inverted items is under vigorous debate in psychometric research. However, especially in the field of the *Dark Tetrad* – a compound of the aversive yet subclinical traits *Machiavellianism*, *narcissism*, *psychopathy*, and *sadism* –, the use of items in which high endorsement indicates low levels appears promising to obtain more information about low scores on the four traits. In this preregistered research ($N = 500$), we developed an alternative version of the *Short Dark Tetrad* (SD4) which – unlike the original SD4 – has a balanced set of regular and inverted items. Following the theory of utilizing inverted items, we demonstrate that more information (in the sense of item response theory) can be obtained from the newly devised *Mixed SD4* (MSD4) as compared to the original SD4. Thereby, the scores of the MSD4 can be validly interpreted in the sense of the underlying traits' theories (i.e., construct validity), and the SD4 and the MSD4 yield highly similar nomological networks. We conclude that including inverted items is advantageous for the assessment of the Dark Tetrad. More generally, we present this case as a demonstration that balanced item sets are necessary to capture traits and behaviors exhaustively.

Keywords: reversed items, scale development, nomological network, aversive traits, factor analysis



Various psychometric scales comprise “regular” items along with inverted items, with a high endorsement of regular (inverted) items indicating high (low) trait levels. Simultaneous use of both item types is supposed to (1) reduce response biases such as acquiescence, (2) ensure respondents’ attention, and (3) extend the range within which information can be obtained by ensuring an exhaustive conceptualization of the constructs. Due to differential polarity, inverted items foster the use of broader retrieval strategies concerning self-related knowledge and thus, account for more differentiated responses. Furthermore, it can be assumed that both regular and inverted items are affected by social desirability, but in opposite directions (e.g., Ray et al., 2016; Weijters & Baumgartner, 2012). Authors including inverted items make use of the decomposition of an observed value into true and error-related aspects (i.e., $X = T + E$). By reducing sources of response

bias, inverted items decrease noise (E), and thus help approximate an individual’s true score, T , through the observed score, X . However, inverted items frequently yield only moderate correlations with regular items, making regular and inverted items usually constitute separate factors rather than a common factor. This obscures factor structures and affects internal consistency. Opponents of inverted items, thus, erroneously conclude that different rather than equal constructs were assessed (cf. Greenberger et al., 2003; Weijters & Baumgartner, 2012, 2022).

From a theoretical perspective, less-than-perfect convergence between regular and inverted items can be desirable because these scores provide complementary rather than redundant information, fostering predictive validity (e.g., Weijters & Baumgartner, 2012). In line with this, regular (inverted) items of the *Rosenberg Self-Esteem Scale* are more strongly related to adaptive (maladaptive) outcomes (Greenberger et al., 2003). One explanation is that items with the same keying direction are more strongly related than items with different keying directions, regardless of whether they assess the same construct or different constructs (van Sonderen et al., 2013). Thus, shared keying

directions pose method effects, account for spurious correlations (Podsakoff et al., 2003), and lower estimates of internal consistencies of mixed scales are artifactual (Weijters & Baumgartner, 2012).

Little Is Known About Low Scores of the Dark Tetrad

We applied the above considerations about the facilitating effects of inverted items to measuring the *Dark Tetrad*. It entails the antagonistic, yet subclinical traits of *Machiavellianism* (manipulation, cynical views, distrust, and striving for power; Christie & Geis, 1970), *Narcissism* (self-exposition, entitlement, and exaggerated self-views; Back, 2018), *Psychopathy* (low conscientiousness, low foresight, ruthlessness, violence, and resistance against authorities; Skeem et al., 2011), and *Sadism* (pleasure-driven enjoyment of violence and of observing others' suffering; Foulkes, 2019). Given that inverted items can help extend the spectra within which information can be obtained, along with the psychometric advantages outlined above, adding inverted items to Dark Tetrad measures appears to be a promising way to enhance assessment. Since the Dark Tetrad is predictive of numerous everyday criteria (see Kowalski et al., 2021, for an extensive overview), the evidence from this study can be of help to general research on aversive behaviors and traits, even those unconcerned with the Dark Tetrad itself.

Utility of Inverted Items in the Dark Tetrad

A contemporary measure of the Dark Tetrad, the *Short Dark Tetrad* (SD4; Paulhus et al., 2021), appears promising in terms of construct validity (Blötner et al., 2022; Paulhus et al., 2021). However, employing *Item Response Theory* the SD4 lacks exhaustive information about individuals low on the four traits (Blötner & Beisemann, 2022). Since inverted items are easier to endorse than regular items in terms of item response theory, they can provide more information about low trait levels (De Ayala, 2022). Unlike the SD4 (Paulhus et al., 2021), its antecessor, the *Short Dark Triad* (Jones & Paulhus, 2014), contains regular and inverted items, but against pertinent recommendations (Weijters & Baumgartner, 2012), the ratio of regular and inverted items is imbalanced. Since the SD4 contains only regular items, we hold that it is more strongly affected by social desirability and other response biases than the Short Dark Triad.

Current Research and Hypotheses

We tested whether an alternative version of the SD4, the *Mixed Short Dark Tetrad* (MSD4) in which half of the items

per subscale are inverted, is comparable to the original scale in terms of construct validity or whether construct validity can even be enhanced. In the MSD4, we tested selected items of the SD4 along with new, inverted items. We examined the original SD4 and the MSD4 in terms of structural properties, the extractability of information across the latent trait spectra, and construct validity. Consistent with earlier research (e.g., Weijters & Baumgartner, 2022), we expected confirmatory factor analyses (CFA) to exhibit less than acceptable fit, which is due to method effects. We hypothesized that it takes method factors to achieve a satisfactory fit for the four subscales and the composite four-factor model. Second, scores of the MSD4 were expected to exhibit more information than the original SD4 scores in terms of item response theory across the four trait spectra. Last, we expected scores of the MSD4 to be at least not inferior to the original SD4 in terms of construct validity. Hence, we hypothesized narcissism scores to correlate positively with extraversion and self-esteem. Machiavellianism scores were expected to correlate positively with cynicism and distrust. Psychopathy scores were assumed to show negative relations with conscientiousness and positive ones with physical aggression and impulsivity. We expected positive relations between sadism scores and physical aggression. Last, since antagonistic and domineering motives underly all Dark Tetrad traits (Paulhus et al., 2021; Semenyna & Honey, 2015), we expected scores of all Dark Tetrad traits to be negatively correlated with agreeableness and positively with dominance seeking.

Method

Sample Size Rationale and Sample Description

We sought to recruit at least 500 participants because this cut-off is advantageous for item response theoretical analyses (Jiang et al., 2016), and exceeds the majority of sample size recommendations for CFA (Kline, 2016). Correlations as high as $|r| = .11$ yield significance at $\alpha = .05$ and $1 - \beta = .80$ (one-sided; Faul et al., 2009). Following our preregistration, we continued recruiting until we had at least 500 respondents who passed two attention checks and an integrity check. The sample comprised 398 women (80%), 100 men (20%), one diverse person, and two did not indicate their gender. Two hundred fifty-three participants were undergraduate students, 205 had a bachelor's degree or above, 32 had less than an academic education, two did not have any degree, and eight reported different education. The age distribution ranged from 18 to 66 years ($M = 34.4$, $SD = 12.1$).

Measures

Dark Tetrad

We used Blötner and colleagues' (2022) German version of the SD4 (original by Paulhus et al., 2021) as a basis for regular Dark Tetrad items. It measures each of the four traits with seven items. We further developed nine items per subscale that reflect descriptions indicating either the opposite pole or the absence of Dark Tetrad traits (as in the Short Dark Triad; Jones & Paulhus, 2014). That is, we aimed to identify polar opposites instead of mere negations of the above characterizations (Weijters & Baumgartner, 2012) and extant theories of the four traits (Back, 2018; Christie & Geis, 1970; Foulkes, 2019; Skeem et al., 2011; see Table S1 in <https://osf.io/2edbr/> for the initial item pool). Simple negations were avoided because they do not necessarily reflect an opposite concept, pose risks of misunderstanding, and cause higher response latency due to cognitive requirements in processing (Swain et al., 2008). All regular and inverted items were presented on the same survey page in randomized order (1 = *strongly disagree* to 5 = *strongly agree*). Inverted items were recoded prior to computations.

Broad Personality

We measured extraversion, agreeableness, and conscientiousness with the same-named 12-item subscales from Danner and colleagues' (2016) German *Big Five Inventory - 2* (1 = *disagree at all* to 5 = *fully agree*). Half of the items per subscale were inverted.

Self-Esteem

We assessed self-esteem with von Collani and Herzberg's (2003) German version of the 10-item *Rosenberg Self-Esteem Scale* (1 = *does not apply at all* to 5 = *applies completely*). Half of the items were inverted.

Cynicism

We assessed cynicism with Blötner and Bergold's (2022) German translation of Leung and Bond's (2004) 20-item *Social Cynicism Scale* (1 = *strongly disbelieve* to 5 = *strongly believe*).

Mistrust and Physical Aggression

To measure physical aggression and mistrust, we utilized the same-named 3-item subscales from Werner and von Collani's (2014) short form of the *German Aggression Scale* (1 = *not applicable* to 4 = *fully applicable*).

Dominance

To measure dominance seeking, we used Suessenbach and colleagues' (2019) same-named 6-item subscale from the *Dominance, Prestige, and Leadership* scale (1 = *does not apply at all* to 6 = *applies perfectly*).

Impulsivity

To measure impulsivity, we used Keye and colleagues' (2009) German 20-item version of the *Urgency Premeditation Perseverance and Sensation Seeking Impulsive Behavior Scale* (1 = *strong disagreement* to 5 = *strong agreement*). In addition to our preregistration, which stated calculating an overall score, we computed scores for *Urgency* (i.e., low self-control due to negative affect), *Lack of Premeditation* (i.e., acting before thinking), *Lack of Perseverance* (i.e., low patience for tedious tasks), and *Sensation Seeking* (i.e., seeking thrill). Two of the five items of the Lack of Perseverance subscale were inverted.

Analysis Plan

Item Selection

In keeping with Ray and colleagues (2016), we employed a sequential approach to select the most eligible items for the new scale. First, we investigated item-total correlations (r_{it} s) among all items assessing the same construct with the same keying direction. We aimed for subscales as short as the original 7-item SD4 subscales. Thus, we selected those four items with the highest r_{it} s per subscale and keying direction. These items were subjected to structural analyses with the R package *lavaan* (version 0.6–8) and the *mean- and variance-adjusted weighted least squares* estimator (Rosseel, 2012). We carried out CFA, *correlated trait-correlated method-1* models (involving a method factor for inverted items), and *correlated trait-correlated method* models (individual method factors for regular and inverted items) both for each subscale and for the composite four-factor model. The models were evaluated using Hu and Bentler's (1999) cut-offs according to which *Comparative Fit Indices* (CFI) > .90, *Root Mean Square Error of Approximation* (RMSEA) < .06, and *Square Root Mean Residuals* (SRMR) < .08 indicate sufficient fit.

Item Response-Theoretical Analyses

The (M)SD4 items were subjected to item response theory analyses with the R package *mirt* (version 1.36.1; Chalmers, 2012), employing the polytomous *graded response model*. We tested local independence through Yen's Q3s, that is, item residual correlations after fitting the model. Local independence was assumed to hold for item-specific |Yen's Q3| < .20 (Chen & Thissen, 1997). We computed *item discrimination indices* (i.e., the degree to which an item distinguishes between persons), *item threshold indices* (the trait level required to endorse a specific response category or higher), *category probability plots* (illustrations of the expected response behavior as a function of the latent trait level), and *reliability plots* (trajectories of reliability across

the spectra) and compared these outputs between the SD4 (Paulhus et al., 2021) and the MSD4. We treated item discriminations as low, moderate, high, and very high when exceeding 0.35, 0.65, 1.35, and 1.70, respectively (Baker, 2001).

Construct Validation

We examined the construct validity of the (M)SD4 subscales by testing the above correlation hypotheses on the nomological networks. To quantify similarities of the subscales purported to measure the same trait, we computed the *Double-Entry Intraclass Correlation* (ICC_{DE}) with the R package *iccde* (version 0.3.5; Blötner & Grosz, 2023) and tested with the R package *diffcor* (version 0.7.2; Blötner, 2022) whether subscales on the same trait correlated differently with the criteria. For the computation of the ICC_{DE} , the correlations of the compared scales of interest with the criteria are appended to each other (i.e., correlations observed for scale 1 to the correlations observed for scale 2 and vice versa). These concatenated vectors are then correlated. In doing so, the profiles' elevations (profile means), scatters (dispersion within the profiles), and shapes (visual form of the trajectories) are aligned. In this way, differences in the profiles cannot be attributed to differences in the distributions. The resultant coefficient can be interpreted as a bivariate correlation (Blötner & Grosz, 2023). This study obtained approval from the Institutional Review Board of the FernUniversität in Hagen and was preregistered. Data, R scripts, and supplements are archived on <https://osf.io/2edbr/>.

Results

Item Selection

For each subscale, we selected those four original SD4 items and those four inverted items with the highest r_{it} s ($.34 \leq r_{it} \leq .56$). As expected, all regular CFA revealed poor fit. Correlated trait-correlated method as well as correlated trait-correlated method-1 models of the narcissism, psychopathy, and sadism subscales yielded a good fit, but correlated trait-correlated method models of the narcissism and psychopathy subscales exhibited Heywood cases and the Dark Tetrad composite model did not converge (see Table 1). However, these issues frequently occur in such models (Fan & Lance, 2017). Thus, the correlated trait-correlated method-1 model should be interpreted. The CFI of the correlated trait-correlated methods-1 model of the Machiavellianism subscale was marginally below Hu and Bentler's (1999) benchmark for acceptable fit, but RMSEA and SRMR indicated good fit (Hu & Bentler, 1999). In this

vein, some scholars argued that common fit conventions are too strict (Heene et al., 2011). This also regards the ostensibly poor fit of the four-factor model. All traits of the Dark Tetrad share antagonistic tendencies, which increases the likelihood of suggested cross-loadings. Unlike fit indices, loadings were satisfactory, which we considered as preliminary evidence for the structure (see Appendix A). The average loadings were consistent across model alternatives. Only a few loadings differed across models, especially those of inverted items, which is frequently the case (e.g., Ray et al., 2016).

Item Response-Theoretical Analyses of Original and Mixed Short Dark Tetrad

Machiavellianism

In the Machiavellianism subscales of the (M)SD4, one of 21 (4 of 28) residual correlations were higher than Chen and Thissen's (1997) cut-off ($|Q3s| > .20$), indicating slight violations of local independence for both subscales. Except for the discrimination indices of the seventh items of the Machiavellianism subscales of the (M)SD4 and the fifth item of the SD4 (α s = 0.38, 0.53, and 0.62), all indices were at least moderate ($\alpha \geq 0.71$). Importantly, the Machiavellianism subscale of the SD4 was on average more difficult to endorse at very low levels (β_1), but the Machiavellianism subscale of the MSD4 was more difficult to endorse at low-to-moderate to high levels (β_2 to β_4 ; see Table 2 and Figure S1 in <https://osf.io/2edbr/>). The comparison of the two subscale alternatives revealed that higher reliability could be obtained for the MSD4 Machiavellianism subscale (see Figure 1).

Narcissism

From 21 (28) residual correlations in the narcissism subscales of the (M)SD4, five (nine) $Q3$ s exceeded Chen and Thissen's (1997) cut-off, suggesting moderate violations of local independence. Except for the seventh item of the SD4 narcissism subscale ($\alpha = .40$), all discrimination indices were at least moderate ($\alpha \geq .76$). It was more difficult to obtain high scores on the SD4 subscale (see Table 2 and Figure S2 in <https://osf.io/2edbr/>). Trajectories of reliability of the two narcissism subscales across the latent spectra were very similar but the MSD4 subscale provided higher reliability at very low scores (i.e., $-4 \leq \theta \leq -2$; Figure 2).

Psychopathy

In the psychopathy subscales of the (M)SD4, 1 of 21 (7 of 28) $Q3$ s exceeded Chen and Thissen's (1997) cut-off, suggesting small (moderate) violations of local independence. All discrimination indices were at least moderate (α s $\geq .70$). On average, all thresholds were lower for the psychopathy subscale of the SD4 than for the respective

Table 1. Different confirmatory factor analyses of the mixed short dark tetrad

Model	$\chi^2(df)$	CFI	RMSEA [90% CI]	SRMR	$M_\lambda (SD_\lambda)$
Machiavellianism					
CFA	87.16 (20)	.75	.08 [.06, .10]	.08	0.44 (0.10)
CTCM-1	45.70 (16) [†]	.89	.06 [.04, .08]	.04	0.40 (0.18)
CTCM	20.68 (12) [†]	.97	.04 [.00, .07]	.03	0.37 (0.16)
Narcissism					
CFA	121.58 (20)	.84	.10 [.08, .12]	.07	0.54 (0.08)
CTCM-1	48.54 (16) [†]	.95	.06 [.04, .08]	.04	0.51 (0.12)
CTCM*	37.47 (12) [†]	.96	.07 [.04, .09]	.03	0.51 (0.11)
Psychopathy					
CFA	178.02 (20)	.64	.13 [.11, .14]	.10	0.45 (0.07)
CTCM-1	39.68 (16) [†]	.95	.05 [.03, .08]	.04	0.41 (0.18)
CTCM*	23.37 (12) [†]	.97	.04 [.02, .07]	.03	0.46 (0.46)
Sadism					
CFA	77.96 (20)	.82	.08 [.06, .09]	.07	0.50 (0.13)
CTCM-1	41.40 (16) [†]	.92	.06 [.04, .08]	.04	0.47 (0.19)
CTCM	21.83 (12) [†]	.97	.04 [.01, .07]	.03	0.48 (0.16)
Four-factor model					
CFA	907.95 (458)	.68	.04 [.04, .05]	.08	0.48 (0.10)
CTCM-1	847.16 (442) [†]	.71	.04 [.04, .05]	.07	0.47 (0.14)
CTCM	No convergence				

Note. CFA = Confirmatory factor analysis; CTCM-1 = Correlated Trait-Correlated Method-1 Model; CTCM = Correlated Trait-Correlated Method Model; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Square Root Mean Residual; M_λ and SD_λ = Mean and standard deviation of loadings. *Heywood cases occurred. [†]Model fit differed at $p < .05$.

MSD4 subscale (see Table 2 and Figure S3 in <https://osf.io/2edbr/>). The psychopathy subscale of the MSD4 yielded higher reliability than the respective SD4 scale, especially for high levels (Figure 3).

Sadism

In the sadism subscales of the (M)SD4, six of 21 (ten of 28) Q3s exceeded Chen and Thissen's (1997) cut-off, indicating moderate violations of local independence. Except for the seventh item of the SD4 sadism subscale ($\alpha_s = .38$), all discrimination indices were at least moderate ($\alpha_s \geq .67$). On average, the SD4 sadism items were easier to endorse than the MSD4 items at low to moderate levels (β_1 to β_3), but at high levels, the SD4 sadism subscale was more difficult (β_4 ; see Table 2, see also Figure S4 in <https://osf.io/2edbr/>). Sufficient reliability (e.g., .60) could be extracted only within narrow ranges (see Figure 4), but this range was broader for the MSD4 subscale ($0 \leq \theta \leq 4$) than for the SD4 subscale ($-1 \leq \theta \leq 4$).

Construct Validity and Similarities Between Original and Mixed Short Dark Tetrad

Machiavellianism

Table 3 provides the correlations, correlation differences, Cronbach's α_s , and agreements of the nomological

networks. Both Machiavellianism subscales correlated positively with cynicism ($r_s = .53$ and $.44$, $p_{\text{difference}} = .002$, all $ps < .001$, if not stated otherwise), mistrust ($r_s = .28$ and $.25$, $p_{\text{difference}} = .35$), and dominance ($r_s = .41$ and $.51$, $p_{\text{difference}} = .001$), and negatively with agreeableness ($r_s = -.20$ and $-.29$, $p_{\text{difference}} = .005$; first [second] coefficients mentioned refer to [M]SD4 subscale). Both subscales were highly correlated ($r = .72$) and yielded very similar nomological networks, $ICC_{DE} = .88$.

Narcissism

The narcissism subscales of the SD4 and the MSD4 correlated positively with extraversion ($r_s = .59$ and $.70$), self-esteem ($r_s = .36$ and $.46$, both $p_{\text{difference}}s < .001$), and dominance striving ($r_s = .39$ and $.31$, $p_{\text{difference}} = .001$), but were unrelated to agreeableness ($r_s = .02$ and $.04$, $ps = 1.00$, $p_{\text{difference}} = .43$). The subscale alternatives were strongly related to each other ($r = .84$) and their nomological networks were highly similar, $ICC_{DE} = .90$.

Psychopathy

The psychopathy subscales of the SD4 and the MSD4 were negatively related to conscientiousness ($r_s = -.27$ and $-.59$, $p_{\text{difference}} < .001$) and agreeableness ($r_s = -.32$ and $-.30$, $p_{\text{difference}} = .49$) and positively related to physical aggression ($r_s = .42$ and $.38$, $p_{\text{difference}} = .15$), dominance striving ($r_s = .44$ and $.37$, $p_{\text{difference}} = .01$), urgency ($r_s = .44$ and $.53$,

Table 2. Item response theoretical parameters of the original and the mixed short dark tetrad

	Regular					Mixed				
	α	β_1	β_2	β_3	β_4	α	β_1	β_2	β_3	β_4
Machiavellianism										
M1	0.71	−5.38	−3.03	−0.71	2.47	1.01	−2.67	−0.46	1.04	3.56
M2	1.16	−1.09	0.92	2.65	4.55	1.02	−3.06	−1.05	0.28	3.43
M3	1.67	−1.94	−0.36	0.75	2.56	1.42	−0.94	0.83	2.32	3.98
M4	1.06	−2.42	−0.43	1.39	3.81	0.81	−2.96	−0.51	1.72	4.76
M5	0.62	−5.71	−3.04	−0.32	4.40	0.90	−0.48	3.29	5.54	6.59
M6	1.28	−2.59	−0.91	0.22	2.90	1.04	−0.59	2.17	3.36	4.67
M7	0.38	−12.93	−8.85	−4.77	1.24	0.53	−4.95	−0.93	1.77	5.52
M8	–	–	–	–	–	1.48	−0.002	1.89	2.61	3.41
Narcissism										
N1	1.31	−1.81	−0.25	1.19	3.28	1.43	−2.07	−0.48	1.14	3.12
N2	1.92	−1.79	−0.40	0.99	2.64	1.31	−1.67	−0.27	1.30	3.26
N3	1.07	−1.48	0.20	2.62	4.62	1.14	0.01	1.36	2.62	4.17
N4	1.62	−1.47	−0.21	1.16	2.84	1.34	−3.37	−1.60	−0.22	2.23
N5	1.45	−3.21	−1.51	−0.19	2.11	1.44	−1.51	0.14	1.21	2.85
N6	1.31	0.02	1.24	2.39	3.77	1.68	−2.55	−1.08	0.09	1.49
N7	0.40	−5.27	−0.23	2.86	9.84	1.23	−3.25	−1.07	0.58	2.63
N8	–	–	–	–	–	0.76	−4.21	−0.79	0.64	3.70
Psychopathy										
P1	1.18	0.49	2.29	3.84	–	1.24	0.38	1.83	2.84	4.69
P2	1.24	−1.38	0.10	1.18	3.27	1.03	−1.57	0.10	1.33	3.76
P3	1.90	0.30	1.44	2.22	3.90	1.12	−1.02	0.98	2.00	4.64
P4	0.79	−2.07	0.65	2.17	5.16	1.33	0.64	2.13	3.51	–
P5	0.98	0.59	2.03	2.38	5.16	1.24	−1.33	1.19	2.47	4.00
P6	1.42	−0.88	0.84	1.71	3.91	0.70	−2.98	0.01	1.50	3.64
P7	0.73	−1.92	0.31	3.52	8.86	0.91	−1.31	0.98	1.92	3.75
P8	–	–	–	–	–	0.96	−2.09	0.61	2.32	5.02
Sadism										
S1	2.81	0.47	1.31	1.85	2.96	2.75	0.37	1.17	2.18	3.05
S2	1.58	0.21	1.05	1.61	2.64	2.57	0.48	1.34	1.91	3.11
S3	0.81	−1.75	−0.01	1.59	4.51	1.53	0.22	1.08	1.65	2.71
S4	3.29	0.36	1.13	2.04	2.84	0.73	−0.63	1.31	2.42	5.43
S5	0.78	−0.60	1.24	2.29	5.10	1.53	0.51	1.69	2.51	3.41
S6	0.67	2.25	3.80	4.70	7.53	0.87	−0.58	3.02	4.38	6.29
S7	0.38	−9.26	−5.43	−2.71	3.74	0.85	−0.36	1.99	2.87	3.96
S8	–	–	–	–	–	1.19	1.26	2.89	3.23	3.92

Note. Regular = Original Short Dark Tetrad as presented by Paulhus et al. (2021); Mixed = Newly devised scale with half the items being inverted; α = Item discrimination; β_{1-4} = Category (transition) specific difficulties. All parameters refer to the polytomous graded response model. Bolded parameters refer to inverted items.

$p_{\text{difference}} = .001$), lack of premeditation ($r_s = .30$ and $.48$), lack of perseverance ($r_s = .21$ and $.48$, both $p_{\text{differences}} < .001$), and sensation seeking ($r_s = .19$ and $.12$, all $p_s < .05$, $p_{\text{difference}} = .02$). Both psychopathy subscale alternatives were strongly correlated ($r = .77$) and their nomological networks were highly similar, $\text{ICC}_{\text{DE}} = .83$.

Sadism

The sadism subscales of the SD4 and the MSD4 were positively related to physical aggression ($r_s = .38$ and $.40$,

respectively) and dominance striving ($r_s = .51$ and $.46$) and negatively related to agreeableness, $r_s = -.34$ and $-.40$, all $p_{\text{differences}} \geq .015$. The sadism subscales of the SD4 and the MSD4 were strongly correlated ($r = .82$) and their nomological networks were almost identical, $\text{ICC}_{\text{DE}} = .97$. Given that the sadism subscales of the (M)SD4 provided information only in relatively narrow areas of the latent continua in the present and in earlier studies alike (e.g., Blötner & Beisemann, 2022), however, these findings should be treated with caution and be replicated.

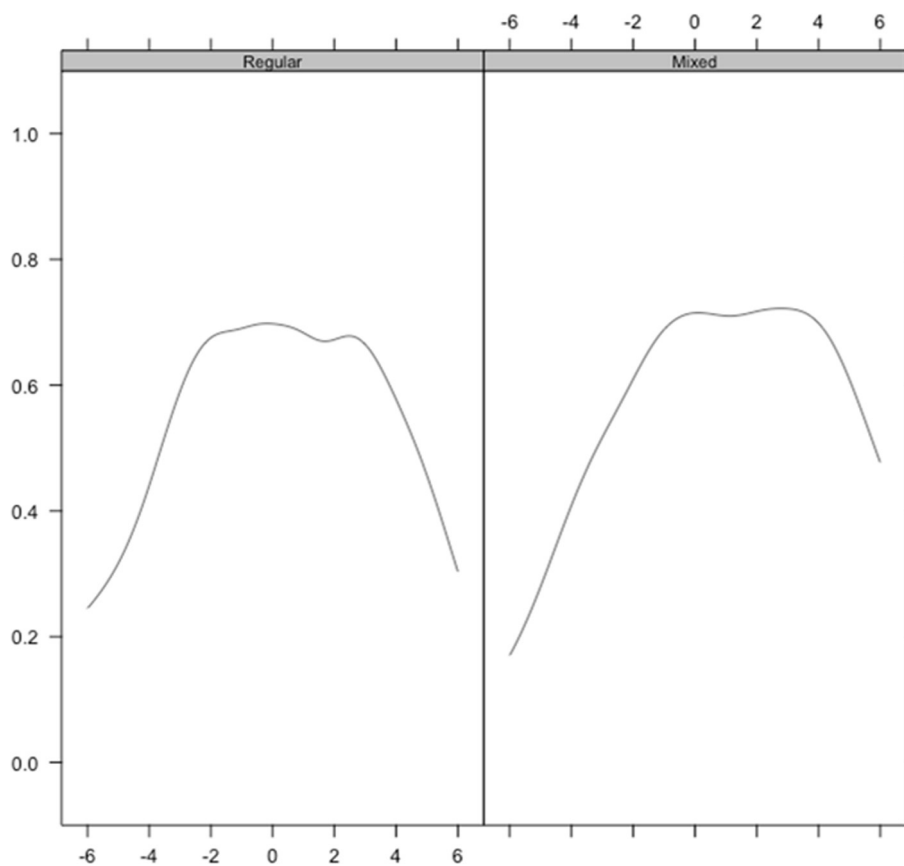


Figure 1. Reliability of the Machiavellianism subscales of the original and mixed short dark tetrad. Reliability is plotted as a function of the latent trait. The left (right) panel displays original (mixed) scales.

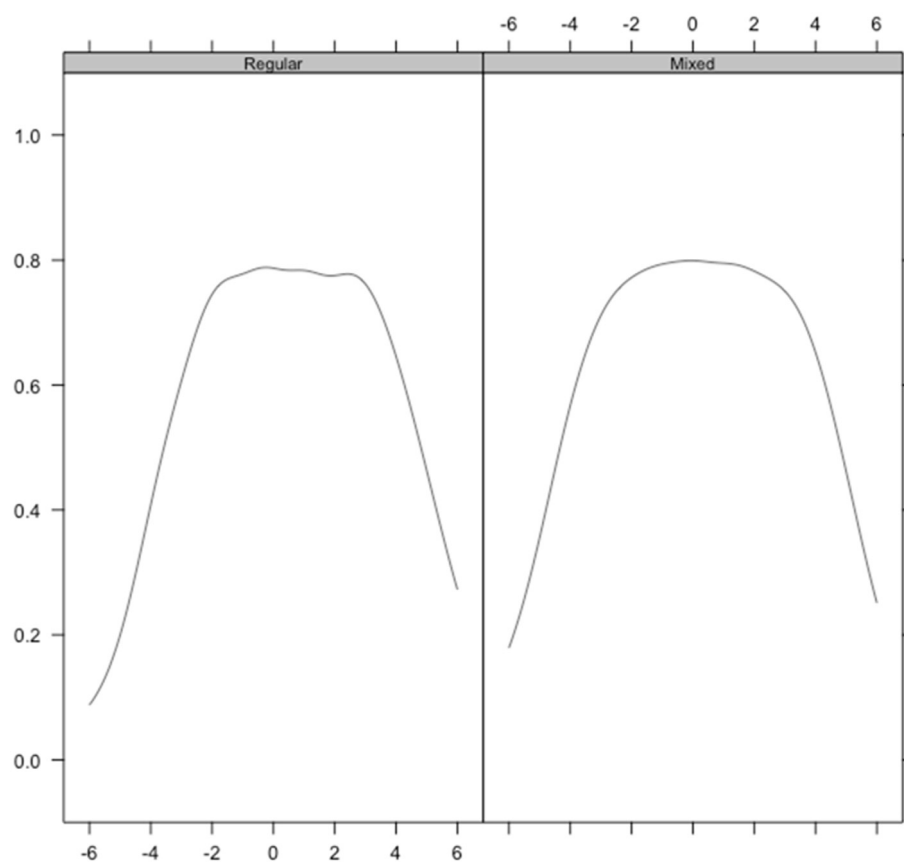


Figure 2. Reliability of the narcissism subscales of the original and mixed short dark tetrad. Reliability is plotted as a function of the latent trait. The left (right) panel displays original (mixed) scales.

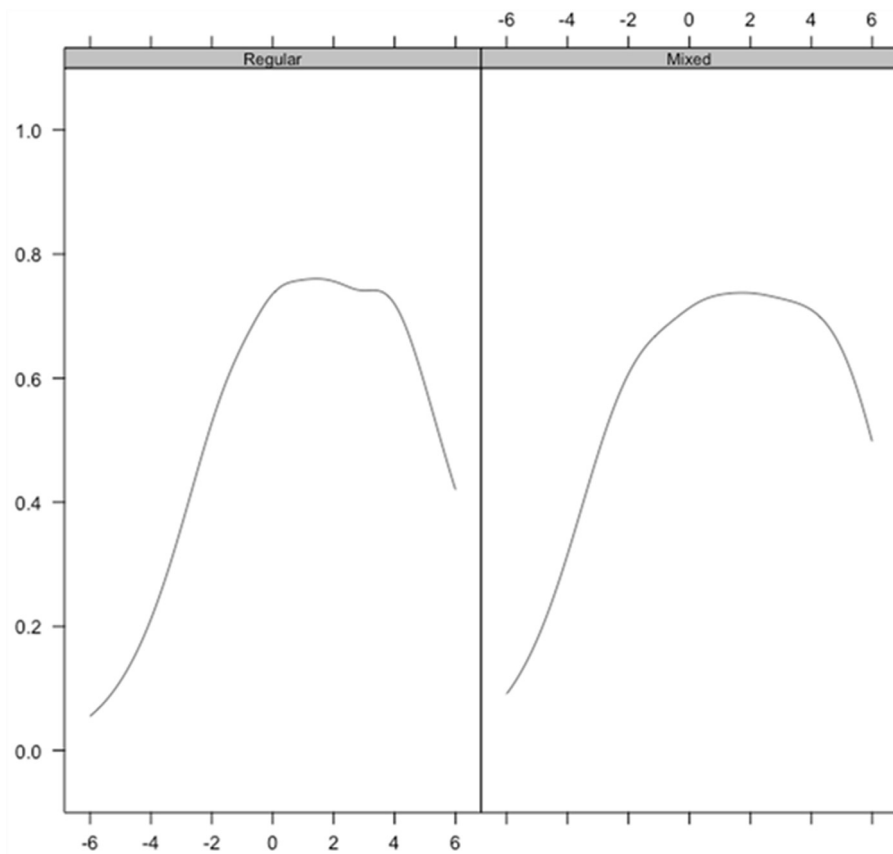


Figure 3. Reliability of the psychopathy subscales of the original and mixed short dark tetrad. Reliability is plotted as a function of the latent trait. The left (right) panel displays original (mixed) scales.

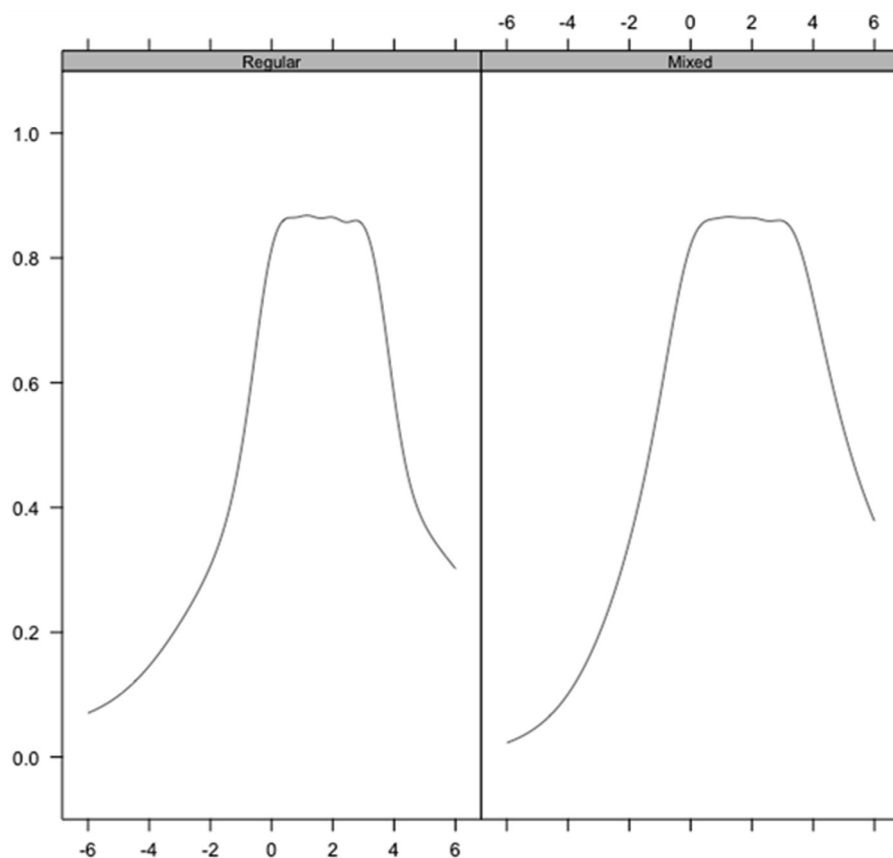


Figure 4. Reliability of the sadism subscales of the original and mixed short dark tetrad. Reliability is plotted as a function of the latent trait. The left (right) panel displays original (mixed) scales.

Table 3. Correlations of the original and the mixed short dark tetrad with the involved constructs

Scale (α)	M _{original}	M _{mixed}	N _{original}	N _{mixed}	P _{original}	P _{mixed}	S _{original}	S _{mixed}
M _{original} (.62)	–	.72	.22	.05	.19	.10	.44	.30
M _{mixed} (.65)	.72	–	.25	.12	.25	.27	.45	.43
N _{original} (.72)	.22	.25	–	.84	.33	.12	.21	.14
N _{mixed} (.77)	.05	.12	.84	–	.28	.10	.10	.07
P _{original} (.66)	.19	.25	.33	.28	–	.77	.37	.31
P _{mixed} (.67)	.10	.27	.12	.10	.77	–	.33	.31
S _{original} (.67)	.44	.45	.21	.10	.37	.33	–	.82
S _{mixed} (.73)	.30	.43	.14	.09	.31	.31	.82	–
Agreeableness (.81)	–.20	–.29	.02	.04	–.32	–.30	–.34	–.40
Extraversion (.88)	–.10	–.06	.59	.70	.19	.06	–.03	–.09
Conscientiousness (.88)	–.08	–.24	.09	.12	–.27	–.59	–.28	–.26
Cynicism (.81)	.53	.44	.07	–.09	.23	.18	.45	.36
Self-esteem (.92)	–.18	–.21	.36	.46	–.10	–.24	–.18	–.19
Mistrust (.73)	.28	.25	–.15	–.23	.20	.21	.30	.23
Physical aggression (.74)	.14	.22	.14	.09	.42	.38	.38	.40
Dominance (.84)	.41	.51	.39	.31	.44	.37	.51	.46
Urgency (.75)	.22	.30	.02	–.11	.44	.53	.24	.19
Lack of premeditation (.80)	–.11	.08	.06	.08	.30	.48	.04	.12
Lack of perseverance (.80)	.11	.26	–.13	–.19	.21	.48	.26	.23
Sensation seeking (.73)	.09	.13	.25	.22	.19	.12	.17	.18
Impulsivity (.81)	.13	.29	.09	.02	.43	.59	.28	.28
ICC _{DE}	.88		.90		.83		.97	

Note. M = Machiavellianism; N = Narcissism; P = Psychopathy; S = Sadism; α = Cronbach's α . Correlations exceeding $|r| = .09$ were significant at $p < .05$. Bolded parameters indicate that the original and the mixed Short Dark Tetrad subscales correlated differently with the criteria, $p < .001$.

Notes on the Contents of the Mixed Items

During the peer-review process, the question arose as to whether the inverted narcissism and psychopathy items rather reflect extraversion and conscientiousness, respectively. Both MSD4 subscale scores correlated more strongly with the stated broad personality scales than the respective SD4 subscale scores (Table 3). We explored how the MSD4 items correlated with extraversion, conscientiousness, and agreeableness scores (see Table S2 in <https://osf.io/2edbr/>). The correlations between the MSD4 items and the stated scores were high in some cases. For instance, the correlation between the overall score of all inverted narcissism items and the overall score of extraversion was $r = .69$. Additionally, the overall score of conscientiousness was correlated at $r = -.59$ and $-.61$ with the inverted psychopathy items “Self-discipline is one of my big strengths.” and “Other people refer to me as being well-structured and good at planning”, respectively, and at $r = -.70$ with the overall score of all inverted psychopathy items (all $ps < .001$). However, (a) extraversion and conscientiousness are key correlates of narcissism and psychopathy, respectively, (b) low agreeableness is supposed to underlie all antagonistic traits, and (c) the contents of the newly devised items are needed to assess the low end of the trait spectra (cf. Kowalski et al., 2021). Additionally, given the favorable properties of the scales in structural,

item-response theoretical, and correlation analyses, we suggest that the items reflect the purported Dark Tetrad traits rather than extraversion, conscientiousness, or agreeableness, respectively. We encourage future research to further examine the conceptual distinctions of the Dark Tetrad, for instance, with other Big Five measures.

Discussion

We developed and probed an alternative version of the Short Dark Tetrad (original by Paulhus et al., 2021) in which half the items of each subscale were inverted, the Mixed Short Dark Tetrad. Our intention to do so was to extend the spectrum of obtainable information in the sense of item response theory. More information, in turn, was expected to benefit construct validity. Particularly, we expected Machiavellianism scores to be positively related to cynicism and mistrust, narcissism scores to be positively related to extraversion and self-esteem, psychopathy scores to be negatively related to conscientiousness and positively to physical aggression and impulsivity, and sadism scores to be positively related to physical aggression. To acknowledge combining elements of all Dark Tetrad traits, we hypothesized scores of all constructs to be related to lower

agreeableness and higher dominance striving. To posit the Mixed Short Dark Tetrad as a feasible alternative to the original Short Dark Tetrad, we expected the new scale to be at least *not* inferior to the original one concerning construct validity. The results were in line with conceptual reasoning and theoretical expectations.

Scale Development and Structural Tests

Structural analyses suggested that regular and inverted items cannot be easily mapped onto the same factor in that regular CFA exhibited (ostensibly) poor fit. This, however, can be attributed to method effects due to item keying (Podsakoff et al., 2003; Weijters & Baumgartner, 2022). Hence, overall fit characteristics were much better if at least one method factor was added. Some correlated trait-correlated method models yielded Heywood cases or did not converge but we did not regard this as a severe drawback because both are common issues in correlated trait-correlated method modeling (Fan & Lance, 2017). However, no version of the overall four-factor model yielded a nominally acceptable fit, which is consistent with the original SD4 and can be explained by shared antagonistic features of all four traits (e.g., Blötner et al., 2022; Kowalski et al., 2021). In light of the conciseness of the four subscales and given that each Dark Tetrad trait has shown to be multidimensional (Kowalski et al., 2021), estimates of internal consistency were generally convincing. However, as opposed to numerous other measures employing inverted items (e.g., Weijters & Baumgartner, 2022), we did *not* find lower estimates of internal consistency for any of the four subscales examined, which underlines the utility of our measure.

Item Response Theoretical Analyses

Ray et al. (2016) tested inverted items to assess *callous-unemotional traits*, which are conceptually similar to the antagonistic traits of the current study. They found regular (inverted) items to discriminate well at high (low) scores on the latent trait spectrum. Our analyses replicated this for a scale of the Dark Tetrad in which half of the items were inverse-coded (see Figures 1–4). The current study also replicated the findings from Blötner and Beisemann's (2022) item response theory analyses of the SD4 in which the sadism subscale exhibited unfavorable properties. A study on the antecedent of the SD4, the Short Dark Triad, demonstrated poor discrimination parameters for inverted items (Dinić et al., 2018). We suggest, however, that distortions of the parameters of the inverted items in the Short Dark Triad are partly due to the imbalanced ratio of regular and inverted items, that is, the method effect of item

polarity is not accounted for. The original SD4 (in which no inverted items occur) also exhibits some poor items in terms of item response theory (see Table 2; see also Blötner & Beisemann, 2022). This being said the discrimination indices of the SD4 and the MSD4 are comparable in many cases, and we conclude that inverted items are not per se an issue. The assumption of local independence was violated for all subscales of the (M)SD4 to different extents, but Chen and Thissen's (1997) cut-off might be too strict for common scales. Furthermore, little is known about the robustness of the analyses against violations of local independence (De Ayala, 2022). As a preliminary conclusion, using the MSD4, we successfully extracted more information for extreme scorers in the Dark Tetrad.

Construct Validation

Our comparisons of the nomological networks of the (M)SD4 demonstrated that the scores of the MSD4 can be interpreted just as validly as the scores of the SD4 because links with criteria were at least equivalent across scales. In several cases, the MSD4 scores were even more strongly related to central criteria than the original SD4 scores (e.g., self-esteem and extraversion in narcissism; conscientiousness and impulsivity in psychopathy). The SD4 and MSD4 scores reflecting the same construct were strongly correlated and yielded highly similar – and in the case of sadism even almost identical – nomological networks. The findings thus advocate in favor of our goal to develop subscales that can be validly interpreted in the sense of the four underlying theories. At the same time, the MSD4 subscales correlated more strongly with different criteria, irrespective of whether the respective measures involved inverted items. Thus, differential method effects not tested in this study might not have played a role in this regard.

The correlations between the SD4 and MSD4 subscales measuring the same trait exceeded 1 once corrected for unreliability. However, since the MSD4 subscales, by and large, outperform the SD4 subscales in item-response theoretical and correlations with other validation criteria, we conclude that the two measures are in no way redundant.

Limitations and Future Directions

Despite promising results concerning structure, scale information, and construct validity, this study was not without limitations. Our sample was skewed in two ways: First, our sample contained more participants self-identifying as women than men, whereby, on average, men report higher scores on antagonistic traits (Kowalski et al., 2021). Second, our study predominantly comprised (undergraduate) students, whereby the endorsement of regular and inverted

items varies as a function of education or verbal comprehension (Gnambs & Schroeders, 2020). In this vein, measurement invariance of the MSD4 across different social groups should be tested to ensure that our results can be generalized. Culture-related invariance is of particular interest because regular and inverted items can be interpreted differently in Western and Eastern cultures (Hamamura et al., 2008). Furthermore, each Dark Tetrad trait is multidimensional, but neither the SD4 nor the MSD4 considers this. Given that short scales are supposed to be used for screening purposes (Kowalski et al., 2021), scholars might wish to build upon the present study. The results showed that inverted items foster psychometric properties and can be central when developing extensive, multifaceted Dark Tetrad scales. Of note, our study was limited to self-reports, raising issues on common method variance.

Conclusion

This study appears to be the first to systematically examine the effects of inverted items in an adapted version of the prominent Short Dark Tetrad (Paulhus et al., 2021). The results suggest that it is worthwhile to also consider low levels of these traits to achieve broader construct coverage. In line with this, Weijters and Baumgartner (2012) suggested that the advantages of using inverted items (higher construct validity) outweigh their caveats (seemingly lower reliability and spurious factors). Indeed, it might even be argued that these caveats arise from misunderstanding the statistical level for its theoretical perspective. The alleged weaknesses of item inversion can also be addressed easily (e.g., method factors or balanced item parceling [Weijters & Baumgartner, 2022] in CFA). Balanced item parceling requires users to intentionally compute parcels with one regular and one inverted item each. In doing so, the respective method effects cancel each other out so that model fit also improves (Weijters & Baumgartner, 2022). Furthermore, the strengths are central to psychometricians from a theoretical perspective. Thus, we encourage scholars to test the utility of inverted items, especially in domains representing malevolent and/or socially undesirable behaviors.

References

Back, M. D. (2018). The narcissistic admiration and rivalry concept. In A. D. Hermann, A. B. Brunell, & J. D. Foster (Eds.), *Handbook of trait narcissism. Key advances, research methods, and controversies* (pp. 57–67). Springer. https://doi.org/10.1007/978-3-319-92171-6_6

Baker, F. B. (2001). *The basics of item response theory* (2nd ed.). ERIC. <http://ericae.net/irt/baker>

Blötner, C. (2022). *diffcor: Fisher's z-tests concerning difference of correlations* (R package version 0.7.2). CRAN. <https://cran.r-project.org/package=diffcor>

Blötner, C., & Beisemann, M. (2022). The Dark Triad is dead, long live the Dark Triad: An item-response theoretical examination of the Short Dark Tetrad. *Personality and Individual Differences*, 199, Article 111858. <https://doi.org/10.1016/j.paid.2022.111858>

Blötner, C., & Bergold, S. (2022). To be fooled or not to be fooled: Approach and avoidance facets of Machiavellianism. *Psychological Assessment*, 34(2), 147–158. <https://doi.org/10.1037/pas0001069>

Blötner, C., & Grosz, M. P. (2023). *iccde: Computation of the double-entry intraclass correlation* (R package version 0.3.5). CRAN. <https://cran.r-project.org/package=iccde>

Blötner, C., & Grüning, D. J. (2023). *An examination of the role of inverted Dark Tetrad items for structural properties and construct validity*. [Open data, materials, and preregistration]. <https://osf.io/2edbr/>

Blötner, C., Ziegler, M., Wehner, C., Back, M. D., & Grosz, M. P. (2022). The nomological network of the Short Dark Tetrad Scale (SD4). *European Journal of Psychological Assessment*, 38(3), 187–197. <https://doi.org/10.1027/1015-5759/a000655>

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, 48(6), 1–29. <https://doi.org/10.18637/jss.v048.i06>

Chen, W.-H., & Thissen, D. (1997). Local dependence indexes for item pairs using item response theory. *Journal of Educational and Behavioral Statistics*, 22(3), 265–289. <https://doi.org/10.2307/1165285>

Christie, R., & Geis, F. L. (1970). *Studies in Machiavellianism*. Academic Press.

Danner, D., Rammstedt, B., Bluemke, M., Treiber, L., Berres, S., Soto, C., & John, O. (2016). Die deutsche Version des Big Five Inventory 2 (BFI-2) [The German version of the Big Five Inventory-2 (BFI-2)]. In *Zusammenstellung sozialwissenschaftlicher Items und Skalen*. <https://doi.org/10.6102/zis247>

De Ayala, R. J. (2022). *The theory and practice of item response theory* (2nd ed.). Guilford Press.

Dinić, B. M., Petrović, B., & Jonason, P. K. (2018). Serbian adaptations of the Dark Triad Dirty Dozen (DTDD) and Short Dark Triad (SD3). *Personality and Individual Differences*, 134, 321–328. <https://doi.org/10.1016/j.paid.2018.06.018>

Fan, Y., & Lance, C. E. (2017). A reformulated correlated trait-correlated method model for multitrait-multimethod data effectively increases convergence and admissibility rates. *Educational and Psychological Measurement*, 77(6), 1048–1063. <https://doi.org/10.1177/0013164416677144>

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>

Foulkes, L. (2019). Sadism: Review of an elusive construct. *Personality and Individual Differences*, 151, Article 109500. <https://doi.org/10.1016/j.paid.2019.07.010>

Gnambs, T., & Schroeders, U. (2020). Cognitive abilities explain wording effects in the Rosenberg Self-Esteem Scale. *Assessment*, 27(2), 404–418. <https://doi.org/10.1177/1073191117746503>

Greenberger, E., Chen, C., Dmitrieva, J., & Farruggia, S. P. (2003). Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: Do they matter? *Personality and Individual Differences*, 35(6), 1241–1254. [https://doi.org/10.1016/S0191-8869\(02\)00331-8](https://doi.org/10.1016/S0191-8869(02)00331-8)

Hamamura, T., Heine, S. J., & Paulhus, D. L. (2008). Cultural differences in response styles: The role of dialectical thinking. *Personality and Individual Differences*, 44(4), 932–942. <https://doi.org/10.1016/j.paid.2007.10.034>

- Heene, M., Hilbert, S., Draxler, C., Ziegler, M., & Bühner, M. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: A cautionary note on the usefulness of cutoff values of fit indices. *Psychological Methods*, 16(3), 319–336. <https://doi.org/10.1037/a0024917>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Jiang, S., Wang, C., & Weiss, D. J. (2016). Sample size requirements for estimation of item parameters in the multidimensional graded response model. *Frontiers in Psychology*, 7, Article 109. <https://doi.org/10.3389/fpsyg.2016.00109>
- 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. <https://doi.org/10.1177/1073191113514105>
- Keye, D., Wilhelm, O., & Oberauer, K. (2009). Structure and correlates of the German version of the Brief UPPS Impulsive Behavior Scales. *European Journal of Psychological Assessment*, 25(3), 175–185. <https://doi.org/10.1027/1015-5759.25.3.175>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Kowalski, C. M., Rogoza, R., Saklofske, D. H., & Schermer, J. A. (2021). Dark Triads, Tetrads, Tents, and Cores: Why navigate (research) the jungle of dark personality models without a compass (criterion)? *Acta Psychologica*, 221, Article 103455. <https://doi.org/10.1016/j.actpsy.2021.103455>
- Leung, K., & Bond, M. H. (2004). Social axioms: A model for social beliefs in multicultural perspective. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 36, pp. 119–197). Elsevier. [https://doi.org/10.1016/S0065-2601\(04\)36003-X](https://doi.org/10.1016/S0065-2601(04)36003-X)
- Paulhus, D. L., Buckels, E. E., Trapnell, P. D., & Jones, D. N. (2021). Screening for dark personalities: The Short Dark Tetrad (SD4). *European Journal of Psychological Assessment*, 37(3), 208–222. <https://doi.org/10.1027/1015-5759/a000602>
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Ray, J. V., Frick, P. J., Thornton, L. C., Steinberg, L., & Cauffman, E. (2016). Positive and negative item wording and its influence on the assessment of callous-unemotional traits. *Psychological Assessment*, 28(4), 394–404. <https://doi.org/10.1037/pas0000183>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Semenyina, S. W., & Honey, P. L. (2015). Dominance styles mediate sex differences in Dark Triad traits. *Personality and Individual Differences*, 83, 37–43. <https://doi.org/10.1016/j.paid.2015.03.046>
- Skeem, J. L., Polaschek, D. L. L., Patrick, C. J., & Lilienfeld, S. O. (2011). Psychopathic personality. *Psychological Science in the Public Interest*, 12(3), 95–162. <https://doi.org/10.1177/1529100611426706>
- Suessenbach, F., Loughnan, S., Schönbrodt, F. D., & Moore, A. B. (2019). The dominance, prestige, and leadership account of social power motives. *European Journal of Personality*, 33(1), 7–33. <https://doi.org/10.1002/per.2184>
- Swain, S. D., Weathers, D., & Niedrich, R. W. (2008). Assessing three sources of misresponse to reversed Likert items. *Journal of Marketing Research*, 45(1), 116–131. <https://doi.org/10.1509/jmkr.45.1.116>
- Van Sonderen, E., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of reverse wording of questionnaire items: Let's learn from cows in the rain. *PLoS One*, 8(7), Article e68967. <https://doi.org/10.1371/journal.pone.0068967>
- Von Collani, G., & Herzberg, P. Y. (2003). Eine revidierte Fassung der deutschsprachigen Skala zum Selbstwertgefühl von Rosenberg [A revised version of the German Rosenberg Self-Esteem Scale]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24(1), 3–7. <https://doi.org/10.1024/0170-1789.24.1.3>
- Weijters, B., & Baumgartner, H. (2012). Misresponse to reversed and negated items in surveys: A review. *Journal of Marketing Research*, 49(5), 737–747. <https://doi.org/10.1509/jmr.11.0368>
- Weijters, B., & Baumgartner, H. (2022). On the use of balanced item parceling to counter acquiescence bias in structural equation models. *Organizational Research Methods*, 25(1), 170–180. <https://doi.org/10.1177/1094428121991909>
- Werner, R., & von Collani, G. (2014). Deutscher Aggressionsfragebogen [German Aggression Scale]. *Zusammenstellung sozialwissenschaftlicher Items und Skalen*. <https://doi.org/10.6102/zis52>

History

Received April 25, 2023

Revision received September 4, 2023

Accepted September 14, 2023

Published online December 13, 2023

EJPA Section / Category Methodological topics in assessment

Conflict of Interest

We have no known conflict of interest to disclose.

Publication Ethics

This study obtained approval from the Institutional Review Board of the FernUniversität in Hagen.

Open Science

We report how we determined our sample size, all data exclusions (if any), all data inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all measures in the study, and all analyses including all tested models. If we use inferential tests, we report exact *p* values and effect sizes. Open Data: The information needed to reproduce all of the reported results is openly accessible. I confirm that there is sufficient information for an independent researcher to reproduce all of the reported results (Blötner & Gruning, 2023).

Open Materials: The information needed to reproduce all of the reported methodology is openly accessible. I confirm that there is sufficient information for an independent researcher to reproduce all of the reported methodology (Blötner & Gruning, 2023).


Preregistration of the Study and Analysis Plan: This study was preregistered with an analysis plan (Blötner & Gruning, 2023). All study materials can be retrieved from <https://osf.io/2edbr/>.

Funding

Open access publication enabled by FernUniversität in Hagen.

ORCID

Christian Blötner

 <https://orcid.org/0000-0001-7415-4756>

David J. Gruning

 <https://orcid.org/0000-0002-9274-5477>

Christian Blötner

FernUniversität in Hagen

Faculty of Psychology

Chair of Personality, Legal Psychology and Assessment

Universitätsstr. 27

58084 Hagen

Germany

christian.bloetner@fernuni-hagen.de

Appendix

Table A1. Item pool and loadings of the Mixed Short Dark Tetrad

Label	Item text	Loadings	
		CFA	CTCM-1
M1	Avoid direct conflict with others because they may be useful in the future. <i>Vermeide direkte Konflikte mit anderen, denn sie könnten in der Zukunft von Nutzen sein.</i>	.36	.41
M2	Flattery is a good way to get people on your side. <i>Schmeicheln ist ein gutes Mittel, um Leute auf deine Seite zu ziehen.</i>	.44	.47
M3	Whatever it takes, you must get the important people on your side. <i>Man muss die wichtigen Personen auf seine Seite ziehen, koste es, was es wolle.</i>	.53	.57
M4	Keep a low profile if you want to get your way. <i>Verhalte dich unauffällig, wenn du deinen Willen durchsetzen willst.</i>	.38	.41
M5i	In my opinion, one should be honest with others. <i>Ich finde, dass man im Umgang mit anderen ehrlich sein sollte.</i>	.41	.33
M6i	You can achieve your goals only by means of honesty. <i>Nur mit Ehrlichkeit kann man seine Ziele erreichen.</i>	.43	.38
M7i	If I do somebody a favor, I do it in order to help, but not to call in a favor myself later. <i>Wenn ich jemandem einen Gefallen tue, mache ich das, um zu helfen und nicht, um selbst später einen Gefallen einfordern zu können.</i>	.34	.26
M8i	I think it is wrong to exploit others for one's goals. <i>Ich halte es für falsch, andere für die eigenen Ziele auszubeuten.</i>	.62	.60
N1	I have a unique talent for persuading people. <i>Ich habe eine einzigartige Begabung, andere zu überzeugen.</i>	.60	.59
N2	I know that I am special because people keep telling me so. <i>Ich weiß, dass ich etwas Besonderes bin, da mir andere das immer wieder sagen.</i>	.58	.58
N3	I'm likely to become a future star in some area. <i>Ich werde in der Zukunft wahrscheinlich ein Star in einem bestimmten Bereich sein.</i>	.55	.56
N4	I have some exceptional qualities. <i>Ich habe einige außergewöhnliche Qualitäten.</i>	.53	.53
N5i	I do not like to be in the center of attention. <i>Ich stehe ungern im Mittelpunkt.</i>	.62	.60
N6i	I regard myself as a follower rather than a leader. <i>Ich halte mich eher für einen Mitläufer als für eine Führungsperson.</i>	.60	.62
N7i	In group activities, I accept the subordinate role. <i>In Gruppenaktivitäten ordne ich mich gerne unter.</i>	.49	.49
N8i	Receiving compliments is generally uncomfortable for me. <i>Komplimente zu bekommen, ist mir im Allgemeinen unangenehm.</i>	.37	.35
P1	I've been in more fights than most people of my age and gender. <i>Ich war in mehr Auseinandersetzungen verwickelt als die meisten Menschen meines Alters und Geschlechts.</i>	.51	.60
P2	I tend to fight against authorities and their rules. <i>Ich neige dazu, gegen Autoritäten und ihre Regeln zu kämpfen.</i>	.41	.48
P3	I sometimes get into dangerous situations. <i>Ich gerate manchmal in gefährliche Situationen.</i>	.51	.58
P4	People often say I'm out of control. <i>Menschen sagen oft, dass ich außer Kontrolle bin.</i>	.50	.51
P5i	I find it important to always abide by rules and laws. <i>Es ist mir wichtig, mich immer an Regeln und Gesetze zu halten.</i>	.59	.51
P6i	Self-discipline is one of my big strengths. <i>Selbstdisziplin ist eine meiner großen Stärken.</i>	.28	.06

(Continued on next page)

Table A1. (Continued)

Label	Item text	Loadings	
		CFA	CTCM-1
P7i	Other people refer to me as being well-structured and good at planning. <i>Andere Menschen sagen von mir, dass ich sehr strukturiert bin und gut planen kann.</i>	.38	.19
P8i	I think thoroughly before I act. <i>Ich denke zuerst sorgfältig nach, bevor ich handle.</i>	.38	.24
S1	I enjoy watching violent sports. <i>Ich genieße es, bei gewalttätigem Sport zuzuschauen.</i>	.64	.68
S2	I really enjoy violent films and video games. <i>Ich genieße gewalttätige Filme und Computerspiele.</i>	.62	.65
S3	Watching a fist-fight excites me. <i>Bei einem Faustkampf zuzuschauen begeistert mich.</i>	.52	.53
S4	Some people deserve to suffer. <i>Manche Menschen verdienen es zu leiden.</i>	.44	.44
S5i	I despise physical violence. <i>Ich verabscheue körperliche Gewalt.</i>	.54	.50
S6i	When I see someone getting hurt, I usually feel compassion. <i>Wenn ich sehe, dass sich jemand verletzt, verspüre ich für gewöhnlich Mitleid.</i>	.42	.38
S7i	I do not think it is okay to insult others verbally. <i>Ich finde es nicht in Ordnung, andere zu beleidigen.</i>	.42	.37
S8i	I find it condemnable to use violence to feel superior to someone else. <i>Ich finde es verwerflich, Gewalt zu verwenden, um sich jemandem überlegen zu fühlen.</i>	.46	.42

Note. M = Machiavellianism; N = Narcissism; P = Psychopathy; S = Sadism; CFA = Confirmatory Factor Analysis; CTCM-1 = Correlated Trait-Correlated Method-1 Model. Suffixed “i” indicates a newly devised, inverted item. All loadings refer to the four-factor model. Regular/original items were adopted from “Screening for Dark Personalities. The Short Dark Tetrad (SD4)”, by D. L. Paulhus, E. E. Buckels, P. D. Trapnell, and D. N. Jones, 2021, *European Journal of Psychological Assessment*, 37(3), 208–222 (<https://doi.org/10.1027/1015-5759/a000602>). © 2021 by Hogrefe. Reprinted with permission (English versions) and from “The Nomological Network of the Short Dark Tetrad Scale (SD4)”, by C. Blötner, M. Ziegler, C. Wehner, M. D. Back, and M. P. Grosz, 2022, *European Journal of Psychological Assessment*, 38(3), 187–197 (<https://doi.org/10.1027/1015-5759/a000655>). © 2022 by the Authors. CC BY-NC 4.0. Reprinted with permission (German version).