



Dark Triad, Dyad, or Core? A Psychometric Evaluation of the Short Dark Triad (SD3) Across Three Countries

Andrew Denovan¹ · Rachel A. Plouffe^{2,3} · Neil Dagnall⁴ · Elena Artamonova⁴ · Christopher Marcin Kowalski⁵ · Donald H. Saklofske⁵

Accepted: 18 November 2024
© The Author(s) 2024

Abstract

Since its introduction to personality psychology literature in 2002, the study of Dark Triad personality traits has gained traction across nations. However, there exists theoretical debate regarding the empirical distinctiveness of traits. Moreover, despite universal study across countries, the Short Dark Triad (SD3) lacks validation for use in all populations. The objective of this study was to scrutinise SD3 performance across three nations, including the United Kingdom ($n = 617$), Canada ($n = 263$), and Russia ($n = 1048$). Specifically, factor structure and item-person functioning of the SD3 was assessed across samples. Exploratory structural equation modelling designated that a three-factor bifactor solution provided superior data-fit. In this model, SD3 items loaded on a general factor, in addition to loading on Machiavellianism, narcissism, and psychopathy dimensions. This enabled scrutiny of the degree to which SD3 items reflected a shared general dimension vs. individual subfactors. Further analyses revealed that the general factor did not possess sufficient variance to disqualify the SD3 as multidimensional. Rasch analyses focusing on the three subscales supported unidimensionality and satisfactory item fit. However, inadequate reliability existed, and items exhibited differential item functioning across nations. Although the SD3 can be considered a valid tool for capturing Dark Triad traits across countries, concerns relating to reliability and DIF suggested that revising SD3 items would enhance measurement precision.

Keywords Cross-cultural · Dark Triad · Machiavellianism · Narcissism · Psychopathy · Validation

Introduction

Why do some individuals adopt insidious behavioural tendencies such as lying, cheating, or engaging in aggression, whereas others do not? This and allied questions motivated academic and clinical study into human malevolence. An

important conceptual landmark was the introduction of the Dark Triad of personality to describe characteristics underlying maladaptive behaviours (see Paulhus & Williams, 2002). The Dark Triad comprises psychopathy, narcissism, and Machiavellianism; these are three interrelated but distinct traits (Paulhus & Williams, 2002). Although narcissism and psychopathy have historically been studied in the context of clinical diagnoses, within personality research, these traits are measured at the subclinical level and exist on a continuum within the general population (Furnham et al., 2013).

Subclinical psychopathy is characterized by several interpersonal and behavioural components, including shallow affect, low empathy, superficial charm, impulsivity, and poor behavioural controls (Southard & Zeigler-Hill, 2016). High narcissism is typified by grandiosity, entitlement, superiority, and self-enhancement (Paulhus & Williams, 2002). Lastly, Machiavellianism, based on the seminal writings of Niccolò Machiavelli, describes individuals who implement manipulative, long-term strategies to exploit others for

✉ Andrew Denovan
a.m.denovan@ljmu.ac.uk

¹ School of Psychology, Liverpool John Moores University, Liverpool L3 3AF, UK

² Department of Psychology, University of Dundee, Dundee, UK

³ Department of Psychiatry, Schulich School of Medicine & Dentistry, University of Western Ontario, London, Canada

⁴ Department of Psychology, Manchester Metropolitan University, Manchester, UK

⁵ Department of Psychology, University of Western Ontario, London, Canada

long-term gain. Persons high in Machiavellianism tend to be cynical of others and disregard conventional morals (Christie & Geis, 1970).

Despite the popularity of research assessing the Dark Triad, there is theoretical debate about whether the traits are empirically distinct. For example, some researchers have suggested that psychopathy and Machiavellianism are redundant (see Kowalski et al., 2021 for a review). McHoskey et al. (1998) were among the first to propose that psychopathy and Machiavellianism, originating from the clinical and personality/social psychology literature respectively, were manifestations of the same construct, but with distinct labels. Similar debates have emerged with reference to overlap between narcissism and psychopathy. At a conceptual level, narcissism shares features with Cleckley's (1976) definition of psychopathy, including superiority, exhibitionism, exploitativeness, and entitlement (Glenn & Sellbom, 2015). This debate, which remains ongoing, extends beyond the conceptual level. Specifically, it remains unclear whether measures of the Dark Triad reliably represent three distinct factors.

Additionally, although the Dark Triad is universally studied across countries, the most commonly used measure, the Short Dark Triad (SD3; Jones & Paulhus, 2014), has not been validated for use with all populations. For example, limited attempts to evaluate the Russian version of the SD3 exist, thus restricting potential for cross-cultural comparisons in the Dark Triad traits. Acknowledging this, the present paper tested competing SD3 factor structures and evaluated the psychometric properties of the scale across national samples (i.e., United Kingdom, Canada, and Russia). This approach has proved an effective method for assessing the psychometric properties of related measures (i.e., Assessment of Sadistic Personality, ASP-8; see Plouffe et al., 2023), in addition to other measures generally (e.g., the Passion Scale; Marsh et al., 2013).

The Dark Triad: measurement considerations

When the Dark Triad was introduced, researchers used extant long form measures to assess subclinical psychopathy, narcissism, and Machiavellianism. These included the Self-Report Psychopathy Scale III (Paulhus et al., 2016), the Narcissistic Personality Inventory (Raskin & Hall, 1979), and the MACH-IV (Christie & Geis, 1970). Combining these instruments produced test batteries that were lengthy, time consuming to complete, and prone to response fatigue. This was especially true when Dark Triad measures were used in conjunction with other assessment tools (Jonason & Webster, 2010).

Consequently, investigators developed specialised short measures in the form of the Dark Triad Dirty Dozen

(DTDD; Jonason & Webster, 2010) and the SD3. The focus of the DTDD was conciseness, using four items per construct to assess the Dark Triad traits. Although the DTDD demonstrated satisfactory reliability and validity, several researchers challenged the construct validity of the scale. Particular concerns included lack of content coverage and high internal consistency coefficients, which indicated scale redundancy (Maples et al., 2014). Subsequently, Jones and Paulhus (2014) developed the 27-item SD3, comprising three 9-item subscales corresponding to each of the triad traits.

Assessment of the psychometric properties of the SD3 by Jones and Paulhus (2014) found that the scale possesses strong self- and informant report reliability and validity, maps well onto traditional long-form Dark Triad measures and is superior to the DTDD in predicting a variety of allied outcome variables (e.g., desire for money and power; Maples et al., 2014). Since its development, the SD3 has become the most widely used measure of the Dark Triad traits, being employed successfully across studies (e.g., Denovan et al., 2021). Despite its popularity, consensus regarding SD3 factor structure has not been achieved.

During development, Jones and Paulhus (2014) used exploratory factor analysis (EFA), followed by exploratory structural equation modelling (ESEM) to assess fit of the proposed three-factor trait structure. The model produced good fit, with only two psychopathy items cross-loading onto Machiavellianism and narcissism (Jones & Paulhus, 2014). Similar findings were reported with Spanish and Serbian translations (Dinić et al., 2018). However, recent investigations posited that SD3 item content may not accurately reflect characteristics underlying Machiavellianism and psychopathy (e.g., Persson et al., 2019). For instance, in an evaluation of SD3 factor structure across samples of United States, British, Canadian, and Australian community members, Persson et al. (2019) found a bifactor model with psychopathy/Machiavellianism and narcissism identified as specific factors best fit the data. The model comprised a general factor, representing an overarching construct (i.e., common "dark" core), and specific factors representing variance that was not subsumed by the general factor (i.e., unique trait variance).

Further evaluations conducted in India and Poland found that bifactor models with narcissism and psychopathy/Machiavellianism included as specific factors (India) and with only psychopathy and Machiavellianism included as separate factors (Poland) produced optimal model fit. These produced better data fit than a three-factor model, unidimensional models with and without correlated errors, and a bifactor Dark Triad model (Rogoza & Cieciuch, 2019; Siddiqi et al., 2020).

To achieve adequate fit, Rogoza and Ciecuch (2019) tested a unidimensional narcissism model and the bifactor psychopathy and Machiavellianism model separately, whereas Siddiqi et al. (2020) sequentially removed items that contributed to model misfit, resulting in an SD3 scale comprising only 12 items. Similar findings emerged in McLarnon and Tarraf's (2017) study in a sample of undergraduate students, with comparisons of various confirmatory factor analytic (CFA) models demonstrating that the bifactor ESEM, in which three specific trait factors were specified and small cross-loadings were permitted, produced the best fit. The structure of a Portuguese version of the SD3 was also evaluated in a sample of at-risk youth and across males and females (Pechorro et al., 2019). Although the original three correlated-factor structure replicated in this sample, the authors removed two problematic items from each of the three trait subscales to reconcile misfit.

Collectively, evidence indicates that across samples and translations, there is no clear consistent SD3 factor structure that represents Dark Triad traits. Although some investigations have shown SD3 items represent distinct constructs both within the original English version (Jones & Paulhus, 2014) and translations (Dinić et al., 2018; Pechorro et al., 2019; Pineda et al., 2020), other research has reported that the measure is best represented by a general latent factor representing a dark or malevolent core, in addition to specific variance attributable to unique elements of narcissism, psychopathy, and Machiavellianism (McLarnon & Tarraf, 2017). Persson et al. (2019) concluded that the dark core subsumed most of the SD3's variance, with few items contributing unique variance beyond this general factor. Given these varied findings, it is crucial that researchers clarify the factor structure of the SD3 across samples and languages to ensure that accurate conclusions are drawn regarding the traits and their associations with relevant attitudes, characteristics, and behaviours.

In addition to clarifying SD3 factor structure, it is imperative that the psychometric equivalence of the measure is established across countries and languages. This standardisation will ensure that cross cultural comparisons are meaningful. To date, the SD3 has been translated and validated across several languages, including Serbian (Dinić et al., 2018), Italian (Somma et al., 2019), and Spanish (Pineda et al., 2020). However, only one conference proceeding has evaluated the validity of a Russian SD3 translation (Egorova et al., 2015). Given the academic interest in investigating associations between Dark Triad traits and attitudes and behavioural outcomes in Russian cultures (e.g., workplace decision-making, Shirokova et al., 2022; time perspective, Jonason et al., 2018), it is important to produce psychometrically sound versions of the SD3. This is particularly true since the SD3 was penned in English and there

are potentially interpretative differences across countries. Hence, it is necessary to examine whether items operate similarly across languages and national samples.

Canada was selected in the present study because the SD3 was developed and validated with North American samples (cf. Jones & Paulhus, 2014). Accordingly, Canada represents a meaningful baseline for comparison with Russia. The UK was chosen because it is comparable with Canada on societal dimensions relevant to the Dark Triad (specifically individualism vs. collectivism) (Wetzel et al., 2020). Both nations are individualistic, whereas Russia is more of a collectivist nation (Hofstede, 2001). Moreover, the SD3 is frequently used in research with UK samples (e.g., Egan et al., 2014; Papageorgiou et al., 2023), therefore it is important to establish whether participants interpret items equivalently to Canada, which should occur due to similarities in language and relevant societal dimensions. Accordingly, findings from this paper will establish the extent to which researchers and clinicians can reliably assess and compare levels of the Dark Triad traits across nations that vary across important dimensions, including cultural values and language. Given that cross-cultural investigations of the Dark Triad are becoming increasingly common (e.g., Jonason et al., 2017, 2020; Valentova et al., 2020), it is important that survey items are interpreted the same way across nations, especially among those countries in which the SD3 has not been extensively validated (i.e., Russia).

An approach well suited to assess item performance is item response theory (IRT) (Persson et al., 2017). IRT refers to a family of psychometric models, which focus on the relationship between item properties, responses from individuals to items, and the underlying trait being assessed. Several studies have used IRT to assess psychometric properties of the SD3 (e.g., Dinić et al., 2018; Persson et al., 2017). Persson et al. (2017) and Dinić et al. (2018) established that psychopathy items were the most challenging to endorse, owing to their relatively greater hostile antisocial content. This was in comparison with Machiavellianism in particular, suggesting that this reflects a more benign construct. Narcissism items were, nevertheless, distributed more evenly across the latent trait. However, studies that use IRT for the SD3 have typically focused on discrete populations/nations (e.g., Serbia: Dinić et al., 2018; United States: Persson et al., 2017). Accordingly, there is a lack of information relating to how the SD3 items perform comparatively across nations.

In this context, we used Rasch modelling, which is an IRT technique that assesses item difficulty and person ability separately. Item difficulty refers to the extent that a latent trait is represented by an item, and person ability denotes an individual's degree of a trait (Amin et al., 2012). Participants with greater 'person ability' are more likely to perform successfully (i.e., endorse more challenging items) than

participants with low ability. Furthermore, Rasch can determine the validity of individual items by examining misfit (i.e., redundancy), threshold ordering relating to response categories, dimensionality, and differential item functioning. Differential item functioning is the degree to which groups interpret items differently. In the current study, this information indicated the extent to which the SD3 provide valid international comparisons.

In summary, achieving consensus on the SD3 factor structure has proven challenging, with debates surrounding the existence of a three-factor model (Dinić et al., 2018; Jones & Paulhus, 2014; Pineda et al., 2020) or various bifactor models with a general “dark” core (McLarnon & Tarraf, 2017; Rogoza & Cieciuch, 2019; Siddiqi et al., 2020). The SD3 factor structures have also been tested inconsistently across samples, with some studies allowing cross-loadings (e.g., using ESEM; McLarnon & Tarraf, 2017) and some removing items contributing to misfit (Pechorro et al., 2019).

In addition, although the SD3 has been translated into many languages (Dinić et al., 2018; Pineda et al., 2020; Somma et al., 2019), cross-national item equivalence of the SD3 has not been empirically examined. This is important to establish, particularly in Russian society, which is distinct from both the Canada and the UK in terms of societal values, attitudes, and language (e.g., Doucerain et al., 2021).

Despite extensive research on the psychometric properties of the SD3, the lack of consensus regarding its factor structure necessitates further investigation and validation across diverse populations. Additionally, although studies have explored the factor structure of its translations, more research is needed to establish psychometric equivalence across languages and nations, especially in regions where the SD3 has not been extensively validated, such as Russia.

Current study

We assessed the psychometric properties (including the structure) of the SD3 across UK, Canadian, and Russian adult samples. Based on past research (Dinić et al., 2018; McLarnon & Tarraf, 2017; Pechorro et al., 2019; Persson et al., 2019; Pineda et al., 2020), the authors tested (1) correlated three-factor, (2) three-factor bifactor, (3) correlated two-factor, and (4) two-factor bifactor ESEM models. The correlated three-factor model examined the degree to which the SD3 was underpinned by latent subfactors of Machiavellianism, narcissism, and psychopathy. The bifactor variant assessed if the SD3 was more effectively underpinned by the three subfactors and a general factor, which captured common variance not accounted for by the three subfactors. The two-factor model tested if two subfactors (i.e., Machiavellianism and psychopathy as one factor, narcissism as

the other) occupied a superior representation of the SD3, whereas the two-factor bifactor scrutinised the extent to which the two subfactors and a general dimension (accounting for additional common/shared variance) represented the factorial composition. Figure 1 includes a schematic representation of these models. Since previous studies have reported cross sample variations in model fit, hypotheses about optimal solutions were not stated. Rather, the research was exploratory with regards to factorial structure.

This study also evaluated the reliability, item hierarchy, and differential item functioning of the SD3 across samples. The English-speaking and Russian SD3 versions have displayed reliability previously (Egorova et al., 2015; Jones & Paulhus, 2014), thus we hypothesized that the items would demonstrate satisfactory reliability across national samples. We also postulated that SD3 items, across national samples, would adequately distinguish between participants with varying levels of each Dark Triad trait. Consistent with prior research, regardless of nationality, psychopathy items were expected to be the most challenging to endorse. In terms of differential item functioning, specific hypotheses were not formulated due to the absence of research comparing the SD3 across nations with IRT techniques.

Collectively, the following research objectives/questions were examined: (1) Do factor analytic results support a correlated three-factor, three-factor bifactor, correlated two-factor, or two-factor bifactor portrayal of the SD3? (2) How does the SD3 perform across nations (specifically the United Kingdom, Canada, and Russia) at the factor and item level?

Materials and methods

Participants

This study recruited three independent national samples using adverts on social networks and a research participant pool (i.e., Canada). Data collation occurred online. Only individuals from specified nations took part. An option in the online survey facilitated this process. This enabled the researchers to filter participants using their IP addresses. Additionally, data collation excluded participants younger than 18 years of age. Recruitment used convenience sampling, obtaining a minimum sample for each nation to limit the possibility of issues with model convergence. The minimum was 200, as recommended for ordinal data by Bandalos (2014) and Forero et al. (2009). Analysis treated data cases possessing z -scores > 3.29 or < -3.29 SDs from the mean as outliers, and accordingly deleted during data screening (i.e., 14 UK, 1 Canada, and 53 Russia). Definitive samples included 617 United Kingdom (UK) respondents,

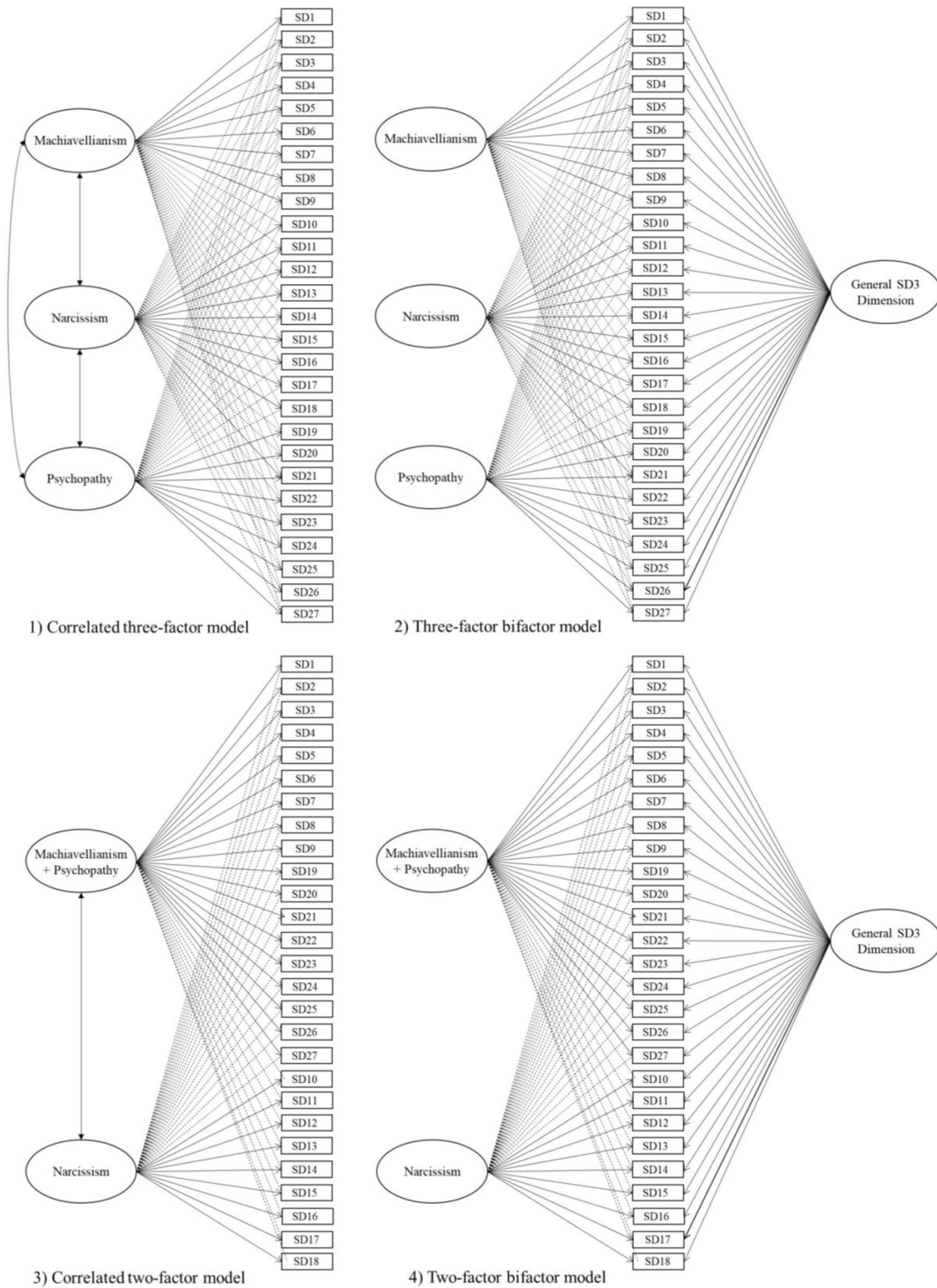


Fig. 1 Schematic representation of competing ESEM solutions for the SD3. Full unidirectional arrows represent factor loadings, dotted unidirectional arrows represent cross-loadings

421 females and 196 males ($M_{\text{age}} = 27.72$, $SD = 11.58$, range = 18–81); 263 Canadian respondents, 138 females and 125 males ($M_{\text{age}} = 18.84$, $SD = 1.84$, range = 18–35); and 1048 Russian respondents, 707 females and 341 males ($M_{\text{age}} = 37.21$, $SD = 10.57$, range = 18–72).

Measure

The Short Dark Triad (SD3; Jones & Paulhus, 2014) measures Machiavellianism, narcissism, and psychopathy using 27 items. Participants recorded their responses on accompanying 5-point Likert type response scales (1 = *disagree strongly* to 5 = *agree strongly*). This study used the SD3 translation by Egorova et al. (2015), which has demonstrated satisfactory reliability, for the Russian sample. In this study, alpha reliability scores for Machiavellianism ranged from 0.68 to 0.84 (i.e., UK $\alpha = 0.70$, Canada $\alpha = 0.84$, and Russia $\alpha = 0.68$); narcissism 0.67–0.72 (i.e., UK $\alpha = 0.72$, Canada $\alpha = 0.67$, and Russia $\alpha = 0.68$); and psychopathy 0.70 to 0.75 (i.e., UK $\alpha = 0.75$, Canada $\alpha = 0.75$, and Russia $\alpha = 0.70$). Reliability results were consistent with previous research, with lower estimates reported for narcissism. In the validation study, Jones and Paulhus (2014) reported an estimate of 0.68 for narcissism. This was similar to those observed in the present study.

Procedure and ethical approval

Interested respondents received information detailing the study aims and objectives. All participants provided informed consent and on doing so received a message comprising a link to the online questionnaire alongside a unique code. Completion of the SD3 was self-paced, and participants completed all items. At the end of the study, participants were debriefed. Ethical review boards based in Western University, Canada, and Manchester Metropolitan University, UK, approved the study.

Analysis

Informed by recent exploratory structural equation modelling (ESEM) applications (e.g., McLarnon & Tarraf, 2017), the authors evaluated the SD3 structure via comparing data fit of correlated three-factor and three-factor bifactor ESEM solutions. Moreover, noting competing research (e.g., Persson et al., 2019) correlated two-factor and two-factor bifactor ESEM iterations (with one factor comprising Machiavellianism and psychopathy, and the other factor representing narcissism) were also considered. Bifactor modelling (Rodriguez et al., 2016) was critical for assessing multidimensionality by comparing systematic item variance relative to an overall factor vs. supplementary sources of

variance (i.e., individual bifactors). ESEM usefully demonstrated the effect of items across factors by allowing cross-loadings. The authors applied target rotation (i.e., assigned zero loadings to items that did not belong to a particular scale relative to the bifactor structure). Parameter estimates and indices of model fit were obtained via weighted least square mean and variance adjusted estimation (WLSMV).

Indices of chi-square, Comparative Fit Index (CFI), Standardized Root-Mean-Square Residual (SRMR), and Root-Mean-Square Error of Approximation (RMSEA) were used to evaluate data-model fit. Good fit criteria denote $CFI \geq 0.90$, $SRMR \leq 0.08$ and $RMSEA \leq 0.08$ (Browne & Cudeck, 1993). Chi-square difference testing was used to compare model fit. Analyses were subsequently re-run using maximum likelihood with robust standard errors (MLR) estimation to facilitate direct comparison using Bayesian Information Criterion (BIC) (Forbes et al., 2015). For interpretation, we utilized the bifactor indices of Rodriguez et al. (2016), including hierarchical omega (ω_h), relative omega (ω), Explained Common Variance (ECV), and Item Explained Common Variance (IECV).

Rasch analysis was used to present supplemental measurement information at the person and item levels. Rasch models can serve “as confirmatory tests of the extent to which scales have been successfully developed according to explicit a priori measurement criteria” (Ludlow et al., 2008, p. 196). The Rasch Rating Scale Model was utilised in this study, focusing on the polytomous variation. Estimation of parameters via Winsteps (Linacre, 2012) used joint maximum likelihood estimation techniques. Akin to prior Rasch validation research (e.g., Denovan et al., 2023), evaluation of the SD3 considered five criteria: rating scale efficacy, reliability, item hierarchy, dimensionality, and differential item functioning.

Rating scale assessment focuses on the utilisation of response categories. The presence of monotonically increasing Rasch-Andrich thresholds and Infit-Outfit statistics ranging from 0.5 to 1.5 are optimal (Bond & Fox, 2015). $Infit-Outfit > 1.5$ and < 2.0 do not bring efficiency to an instrument, but do not degrade measurement either (unless many instances occur) (Wright & Linacre, 1994). Indices of item separation/reliability and person separation/reliability were used to determine reliability. Separation statistics specify item/participant spread along the ability continuum in addition to the amount of distinct item/person ability levels existing within data (Bond & Fox, 2015). Values > 1.5 indicate that the items/sample can be separated into at least two levels, such as high and low difficulty/ability. For reliability, a threshold of 0.7 exists.

Infit-Outfit Mean Square Error (MNSQ) reflected how adequately data conformed to the Rasch model, with MNSQ indicating the discrepancy between observed vs. predicted

values. The same threshold is used as for rating scale efficacy, and misfitting items suggest that these are assessing something different than other items on the measure. Indeed, an adequate instrument should contain items of varying difficulty to distinguish persons on the continuum of ability. Person-item (Wright) maps usefully illustrated targeting of item difficulty to person ability.

Unidimensionality assessment used Principal Components Analysis (PCA) of the residuals. This approach is typically used and tests a core Rasch assumption that a measurement model is unidimensional by establishing if any patterns indicative of additional components exists while controlling for the latent dimension (Han, 2022). Existence of additional components are determined if the first contrast contains an eigenvalue > 2 and accounts for > 15% of variance (Linacre, 2012). Lastly, differential item functioning (DIF) evaluated item equivalence across nations. A significant Mantel-Haenszel *p* value alongside a contrast > 0.50 suggests the presence of bias; specifically, that nations differ in their interpretation of the item/scale (Boone et al., 2014).

Results

Descriptive statistics

Examination of univariate skewness identified no item issues across samples, as values were between -3.0 and +3.0. Multivariate kurtosis (*b2p*) and skewness (*b1p*) via Mardia's test indicated significant deviation from normality: UK sample *b2p*=28.09, *p*<.001, *b1p*=64.65, *p*<.001;

Canadian sample *b2p*=12.36, *p*<.001, *b1p*=114.80, *p*<.001; and Russian sample *b2p*=44.58, *p*<.001, *b1p*=42.64, *p*<.001. However, WLSMV and MLR (used in this study) function effectively with non-normal data (Finney & DiStefano, 2006).

ESEM analyses

Results (computed via Mplus 8.1; Muthén & Muthén, 2018) indicated that the two-factor bifactor model provided a significantly better fit than the two-factor correlated model (UK χ^2 difference = 261.48, *df*=25, *p*<.001; Canada χ^2 difference = 221.13, *df*=25, *p*<.001; Russia χ^2 difference = 764.28, *df*=25, *p*<.001). Moreover, the three-factor bifactor model demonstrated significantly better data-model fit than the three-factor correlated model (UK χ^2 difference = 156.36, *df*=24, *p*<.001; Canada χ^2 difference = 105.99, *df*=24, *p*<.001; Russia χ^2 difference = 296.80, *df*=24, *p*<.001). Using the same parameters, model fit of the three-factor correlated and two-factor bifactor model was identical. This indicated that the three-factor bifactor was superior in comparison with the two-factor bifactor solution. Fit indices for the three-factor bifactor were good for all samples (Table 1).

Factor loadings suggested that Machiavellianism, narcissism, and psychopathy were reasonably well-defined across nations. Relatively weaker loadings existed for Machiavellianism in relation to the Russian sample (UK $\lambda_{\text{Machiavellianism}}M=0.30$; $\lambda_{\text{narcissism}}M=0.43$, $\lambda_{\text{psychopathy}}M=0.38$; Canada $\lambda_{\text{Machiavellianism}}M=0.51$; $\lambda_{\text{narcissism}}M=0.44$, $\lambda_{\text{psychopathy}}M=0.41$; Russia $\lambda_{\text{Machiavellianism}}M=0.11$; $\lambda_{\text{narcissism}}M=0.46$,

Table 1 Fit indices for country-specific SD3 models

Model	χ^2 ^a	df ^a	CFI ^a	SRMR ^a	RMSEA (90% CI) ^a	BIC ^b
Two-factor correlated						
UK	1187.52**	298	0.87	0.05	0.07 (0.06–0.07)	48986.89
Canada	876.30**	298	0.81	0.07	0.08 (0.07–0.09)	19990.14
Russia	2440.09**	298	0.77	0.06	0.08 (0.08–0.09)	74584.05
Two-factor bifactor						
UK	876.12**	273	0.91	0.04	0.06 (0.05–0.06)	48912.73
Canada	592.57**	273	0.90	0.05	0.06 (0.05–0.07)	19915.40
Russia	1198.36**	273	0.90	0.03	0.05 (0.05–0.06)	74093.37
Three-factor correlated						
UK	876.12**	273	0.91	0.04	0.06 (0.05–0.06)	48912.73
Canada	592.57**	273	0.90	0.05	0.06 (0.05–0.07)	19915.40
Russia	1198.36**	273	0.90	0.03	0.05 (0.05–0.06)	74093.37
Three-factor bifactor						
UK	723.20**	249	0.93	0.03	0.05 (0.05–0.06)	48915.42
Canada	483.24**	249	0.92	0.04	0.06 (0.05–0.06)	19948.04
Russia	832.60**	249	0.94	0.03	0.04 (0.04–0.05)	73960.74

** χ^2 significant at *p*<.001. CFI=comparative fit index; SRMR=standardized root-mean-square residual; RMSEA=root-mean-square error of approximation; BIC=Bayesian Information Criterion. (a) Estimated using weighted least square mean and variance adjusted estimation; (b) Estimated using maximum likelihood-robust estimation

$\lambda_{\text{psychopathy}} M=0.43$). Slightly weaker average loadings existed for a general factor in all samples (UK $\lambda M=0.36$, Russia $\lambda M=0.30$, Canada $\lambda M=0.32$).

Hierarchical omega (ω_h) for the general factor was below the threshold of 0.8 (i.e., UK=0.65, Canada=0.54, Russia=0.55). Moreover, ECV supported multidimensionality (UK=0.54, Canada=0.42, Russia=0.47), as did IECV. Only 18% of UK items, 14% of Canada items, and 18% of Russia items possessed IECV > 0.8. Closer inspection of items with generally high IECV (i.e., Machiavellianism item 8 [SD3_8], narcissism item 9 [SD3_18], and psychopathy items 6 [SD3_24] and 9 [SD3_27]) indicated shared features of self-centredness. Moreover, higher IECV typically existed for Machiavellianism and psychopathy items. Relative omega was > 0.20 for all subscales aside from Machiavellianism in the Russian sample. Taken together, these results indicated that the general factor did not possess a sufficient quantity of variance independent of the bifactors. Accordingly, Rasch analysis focused on the subscale level.

Rasch analysis

Assessment of response options across subscales (Table 2) suggested that these were utilised in a way that was supportive of scale construction. For instance, item difficulty and respondent ability increased with each response category increase from *disagree strongly* to *agree strongly* (Machiavellianism -0.94 to 1.32 , narcissism -1.33 to 0.77 , psychopathy -1.64 to 0.03). Outfit MNSQs for each category fell below 2.0, reflecting expected category usage (Linacre, 2012). However, for all subscales, fewer participants than expected used the response category *neither agree nor*

disagree, as denoted by the coherence statistic < 40%. Thus, there was a preference for definitive responses.

Scrutiny of summary fit statistics for each nation revealed low person reliability for some of the subscales (i.e., Russia Machiavellianism 0.57, psychopathy 0.57; Canada narcissism 0.65) alongside Person Separation Indices (PSIs) close to 1. This indicated that participants were not being separated into more than one ability level and suggested either a need for more items, or for a broader sample of people ability (Linacre, 2012). Remaining subscales exhibited marginal-to-acceptable person reliability (UK Machiavellianism 0.69, narcissism 0.69, psychopathy 0.67; Russia narcissism 0.69; Canada Machiavellianism 0.83, psychopathy 0.70). Item reliability, > 0.9, was high for all subscales.

Assessment of item fit statistics (Table 3) revealed that all items across the subscales possessed adequate Infit and Outfit MNSQs between 0.5 and 2.0, alongside positive and fairly strong Point Measure Correlations (i.e., ≥ 0.40). Wright maps (Figs. 2, 3 and 4) illustrated how items were arranged along the latent traits of Machiavellianism, narcissism and psychopathy at the total sample and nation level. Machiavellianism items 1 [SD3_1] and 7 [SD3_7] were consistently easy to endorse across all samples, whereas items 2 [SD3_2], 5 [SD3_5], and 6 [SD3_6] were difficult items. For the Canadian sample, mean endorsement was slightly greater than the average item difficulty. However, the UK sample exhibited mean endorsement parallel to mean item difficulty and mean endorsement was lower than mean item difficulty for the Russian sample. This indicates that the English language Machiavellianism subscale ranged from being slightly easy to complete to possessing adequate sample targeting, whereas the Russian version was slightly difficult to complete.

Table 2 Rating scale effectiveness

	Category	Count (%age)	Avg. Measure	Infit MNSQ	Outfit MNSQ	Coherence
Machiavellianism	1 Disagree strongly	1371 (8)	-0.94	1.01	1.09	79%
	2 Disagree	3613 (21)	-0.39	1.05	1.09	43%
	3 Neither agree nor disagree	4199 (24)	0.01	0.90	0.91	34%
	4 Agree	6130 (35)	0.61	0.93	0.94	55%
	5 Agree strongly	2039 (12)	1.32	1.05	1.04	68%
Narcissism	1 Disagree strongly	2031 (12)	-1.33	0.96	0.98	76%
	2 Disagree	5561 (32)	-0.66	0.97	0.97	50%
	3 Neither agree nor disagree	4561 (26)	-0.17	0.91	0.92	35%
	4 Agree	4281 (25)	0.28	1.01	1.03	53%
	5 Agree strongly	918 (5)	0.77	1.16	1.18	66%
Psychopathy	1 Disagree strongly	4978 (29)	-1.64	0.94	0.97	78%
	2 Disagree	6575 (38)	-0.91	0.96	0.90	47%
	3 Neither agree nor disagree	3022 (17)	-0.49	0.89	0.90	31%
	4 Agree	2310 (13)	-0.22	1.07	1.16	47%
	5 Agree strongly	467 (3)	0.03	1.24	1.45	63%

MNSQ mean square statistics. Total count (%age) = total number of endorsements for a response category and the percentage of the overall total that this represents

Table 3 Item Fit statistics

Item	Difficulty	Infit MNSQ	Outfit MNSQ	PTMEA Corr.
Machiavellianism				
SD3_1	-0.63	1.32	1.38	0.40
SD3_2	0.69	1.08	1.11	0.50
SD3_3	0.31	0.79	0.83	0.58
SD3_4	-0.12	0.94	0.98	0.54
SD3_5	0.41	0.87	0.86	0.68
SD3_6	0.42	0.93	0.94	0.66
SD3_7	-1.19	1.00	0.97	0.51
SD3_8	0.38	0.95	0.98	0.55
SD3_9	-0.26	1.15	1.15	0.48
Narcissism				
SD3_10	-0.16	0.84	0.86	0.59
SD3_11	-0.10	0.94	0.95	0.60
SD3_12	0.36	0.77	0.79	0.58
SD3_13	0.40	0.82	0.82	0.63
SD3_14	-0.60	1.00	1.02	0.53
SD3_15	0.15	1.25	1.28	0.44
SD3_16	0.37	1.09	1.10	0.56
SD3_17	0.52	1.23	1.25	0.46
SD3_18	-0.95	1.07	1.06	0.46
Psychopathy				
SD3_19	0.40	0.85	0.88	0.54
SD3_20	-0.24	1.15	1.30	0.36
SD3_21	0.07	0.92	0.93	0.56
SD3_22	0.28	0.92	0.93	0.55
SD3_23	-0.59	1.03	1.08	0.55
SD3_24	-0.43	0.83	0.87	0.60
SD3_25	0.17	1.56	1.58	0.43
SD3_26	0.28	1.05	1.05	0.54
SD3_27	0.07	0.81	0.81	0.59

MNSQ mean square statistics; PTMEA point measure correlation. The MNSQ acceptable limits to productive measurement were 0.5 to 2.0. Values beyond these limits are considered misfitting

For narcissism, items 5 [SD3_14] and 9 [SD3_18] were consistently easy to endorse, while items 3 [SD3_12], 4 [SD3_13], and 8 [SD3_17] were consistently difficult to endorse. Across nations, mean endorsement appeared to be lower than average item difficulty, signifying that the narcissism subscale was difficult to endorse. Psychopathy items 3 [SD3_21] and 4 [SD3_22] were consistently difficult to endorse and item 5 [SD3_23] was the easiest to endorse across nations. Mean endorsement was considerably lower than mean item difficulty in all samples apart from Russia, demonstrating that this subscale was difficult to complete.

PCA of the residuals assessed dimensionality for each subscale. For Machiavellianism, a single Rasch dimension explained 41.1% of variance. Unexplained variance in the first contrast was 10.6% alongside an Eigenvalue of 1.6. The narcissism subscale evidenced 37.5% of explained variance with one dimension. The first contrast accounted for 11.5% variance and an Eigenvalue of 1.7. A single dimension for

psychopathy explained 33.9% of variance with a first contrast accounting for 11.9% alongside an Eigenvalue of 1.6. This is sufficient evidence in support of unidimensionality (Linacre, 2012).

DIF testing (Fig. 5) revealed no significant DIF across nations for Machiavellianism. For narcissism, significant DIF existed for items 8 [SD3_17] and 9 [SD3_18]. Moreover, significant DIF existed for psychopathy items 1 [SD3_19], 3 [SD3_21], 5 [SD3_23], and 6 [SD3_24]. In all instances, significant DIF was evident for the Russian sample in comparison with the Canadian and UK samples. No meaningful DIF was observed when comparing UK and Canadian samples.

Discussion

This paper evaluated the psychometric characteristics of the SD3 in adult samples from the UK, Canada, and Russia (i.e., factor structure, reliability, item hierarchy, and DIF). Four ESEM models were tested: (1) correlated three-factor, (2) three-factor bifactor, (3) correlated two-factor, and (4) two-factor bifactor models. Analysis found that the two- and three-factor bifactor models were superior to the two-factor correlated model. Given the number of parameters, fit indices for the two-factor bifactor and three-factor correlated model were identical.

Superior fit for the three-factor bifactor solution suggested that the SD3 includes a general dimension, in addition to Machiavellianism, narcissism, and psychopathy, that contributes variance to item responses. This supports the supposition of McLarnon and Tarraf (2017) suggesting that SD3 responses indicate the presence of an overarching factor, which captures general malevolent characteristics, alongside the influence of specific Dark Triad factors. Indeed, items driving the general factor included those reflecting self-centredness or exploitation for personal gain (i.e., “*Make sure your plans benefit yourself, not others*”; “*I’ll say anything to get what I want*”), which reflect (dis)agreeable and dishonesty features proposed to represent overlap in the Dark Triad (Hodson et al., 2018).

Although the three-factor bifactor model showed superior fit, the ECV and IECV values suggested that a general “dark” personality factor lacked a significant amount of variability from the bifactors. Accordingly, mean general “dark” factor loadings were typically lower than the bifactors, ranging from 0.30 (Russia) to 0.36 (UK). Therefore, some non-redundant variance existed for the general factor, but this was not substantial enough to disqualify the SD3 as primarily multidimensional.

In terms of model fit, factor analytic results were in line with those reported by McLarnon (2022) (i.e., reported

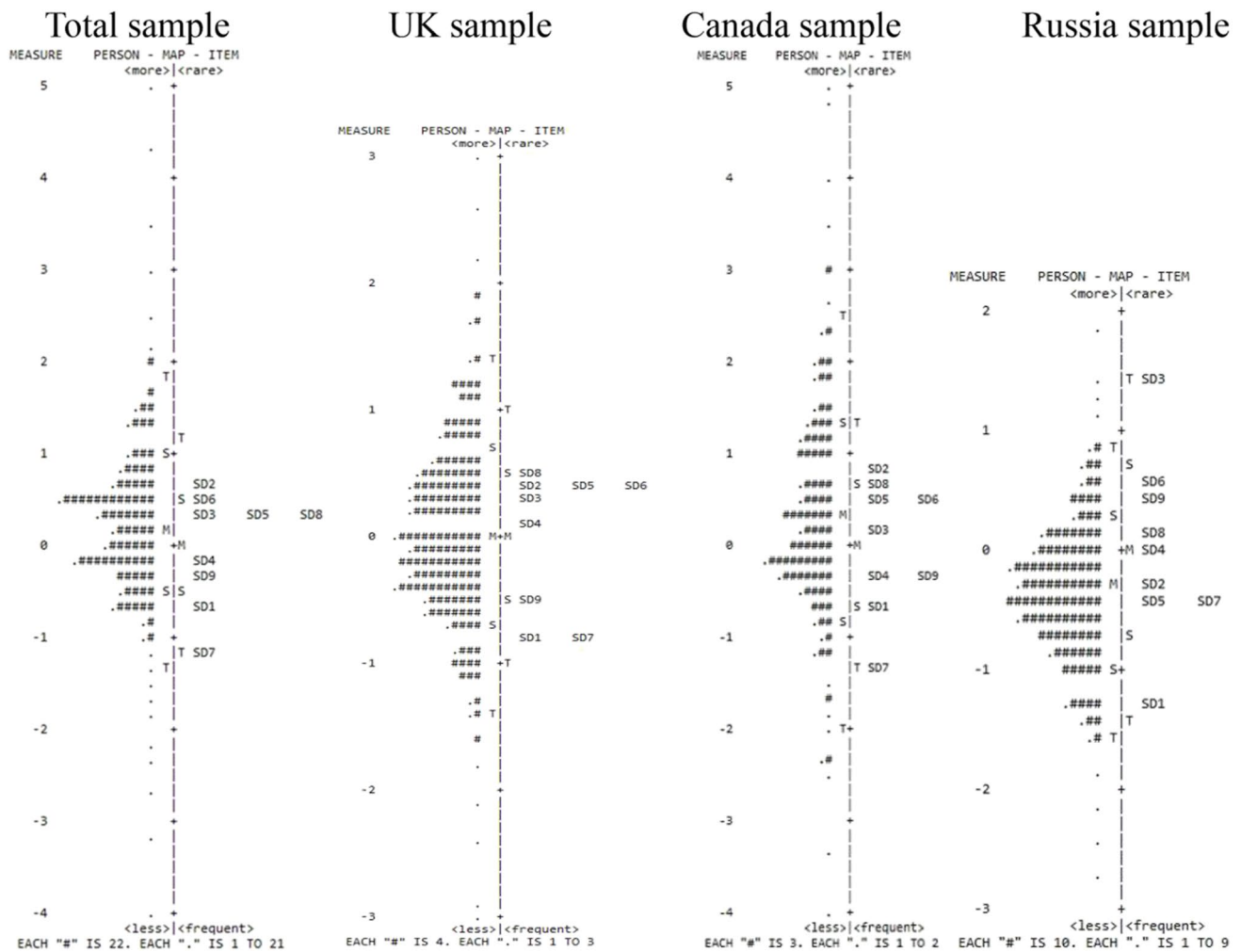


Fig. 2 Person-Item Maps of Machiavellianism Across Nations. The participants are on the left of each dashed line and more able participants are located at the top of each map. Items are located on the right

of each dashed line and more difficult items are also located at the top of each map. *Note.* M=Mean persons' ability or mean items' difficulty; S=one standard deviation; T=two standard deviations

CFI=0.94, RMSEA=0.05, SRMR=0.02). With regards to Persson et al. (2019), findings did not support a two-factor conceptualisation. This could be due to the use of model estimation techniques. Persson et al. (2019) utilised CFA, and it is possible that the association between Machiavellianism and psychopathy (0.90) was increased due to the restrictive nature of this approach. Indeed, Machiavellianism and psychopathy associations in this study for the correlated two-factor (UK=0.42, Canada=0.17, Russia=0.24) and three-factor iterations (UK=0.37, Canada=0.32, Russia=0.23) were weaker, suggesting a clearer differentiation of constructs.

Moreover, Machiavellianism and psychopathy items likely possessed multiple sources of variance in the Persson et al. (2019) study (owing to the superior bifactor structure). This includes variance associated with the focal construct and competing sources (e.g., variance affiliated with conceptually related constructs and a global dimension) (Morin

et al., 2016). However, CFA does not permit scrutiny of item variance in relation to conceptually similar constructs as efficiently as ESEM (Morin et al., 2013). Again, this restriction may have impacted model fit, whereas in the current study use of ESEM facilitated the observation that the SD3 possesses non-redundant variance linked with focal constructs, related constructs (e.g., Machiavellianism with psychopathy), and a general dimension (albeit not substantial).

However, consistent with Persson et al. (2017), higher common variance existed for Machiavellianism and psychopathy. For example, “*Make sure your plans benefit yourself, not others,*” and “*People who mess with me always regret it*”. Though Machiavellianism and psychopathy are conceptually distinct (Kowalski et al., 2021), individuals high in psychopathy are more likely to engage in impulsive sensation-seeking. Additionally, those high in Machiavellianism are more likely to engage in long term, strategic manipulation. Thus, it is possible that the jangle fallacy,

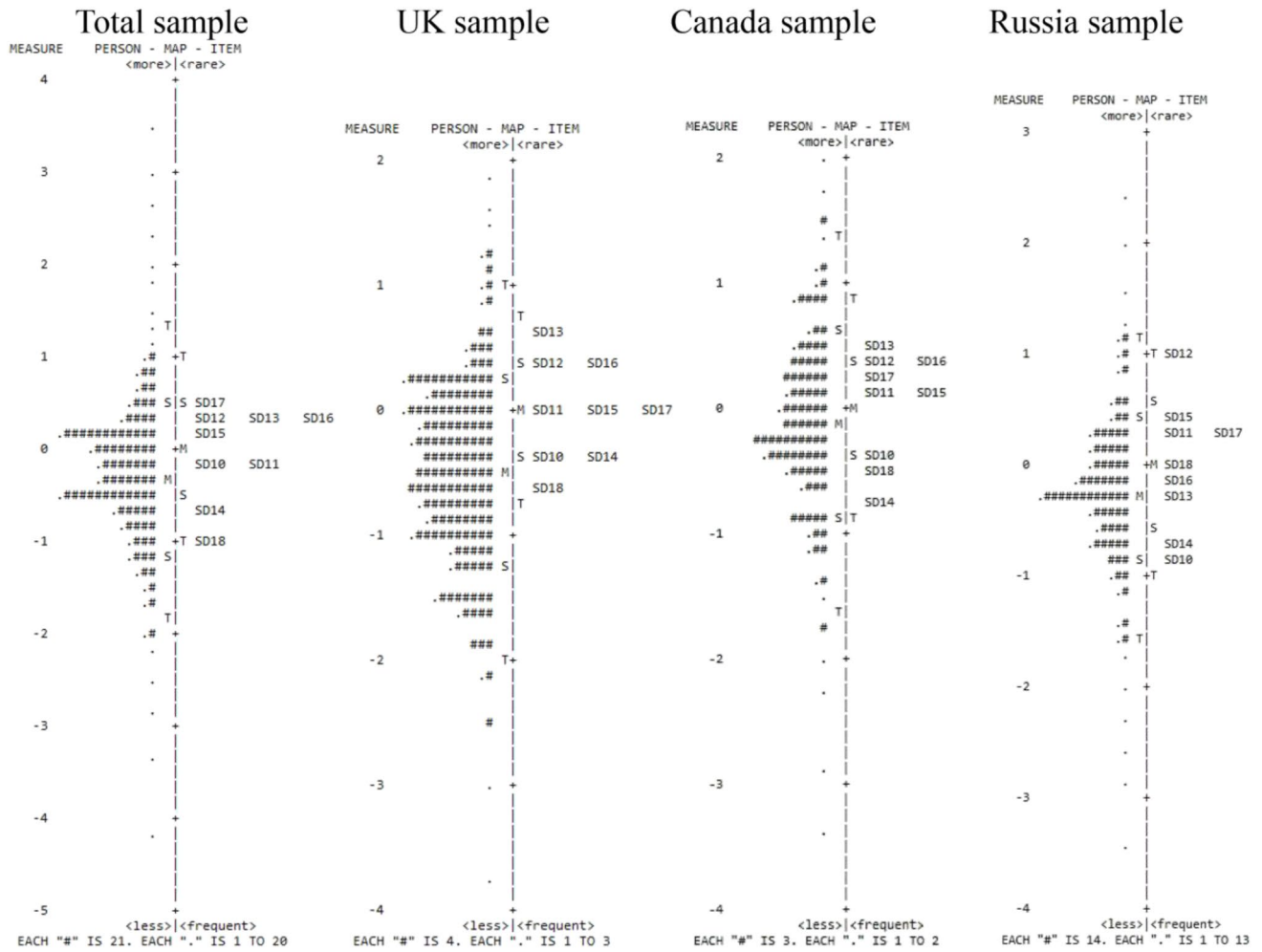


Fig. 3 Person-Item Maps of Narcissism Across Nations. The participants are on the left of each dashed line and more able participants are located at the top of each map. Items are located on the right of

each dashed line and more difficult items are also located at the top of each map. *Note.* M=Mean persons' ability or mean items' difficulty; S=one standard deviation; T=two standard deviations

whereby items reflect content that does not effectively differentiate between the two constructs, accounted for shared variance between SD3 psychopathy and Machiavellianism items (e.g., McHoskey et al., 1998).

This notion concurs with Rogoza and Ciecuch (2019), who using specific factors measuring sensation-seeking driven by impulsivity (psychopathy) and impulse regulation (Machiavellianism), concluded that SD3 Machiavellianism and psychopathy items were only (weakly) differentiated. This aligned with the present study, where one general “dark” personality factor did not adequately explain common variance across items. Acknowledging this and other studies assessing the fit of various two- and three-factor (bifactor and correlated factor) models (Persson et al., 2019; Rogoza & Ciecuch, 2019; Siddiqi et al., 2020), it is evident that the SD3 should be enhanced to improve construct validity and more precisely discriminate between Machiavellianism and psychopathy.

In addition to investigating the structure of the SD3 across countries, this paper used Rasch analysis to evaluate item fit statistics across subscales. Analysis revealed that infit and outfit MNSQs and Point Measure Correlations (PTMEAs) were adequate. Specifically, the MNSQs ranging between 0.5 and 2.0 indicated that items did not exhibit too much or too little variation between expected and observed scores (Bond & Fox, 2015; Wright & Linacre, 1994). Therefore, SD3 responses were generally consistent with expected response patterns. Similarly, PTMEAs ≥ 0.40 specified that SD3 items adequately discriminated between respondents (Bond & Fox, 2015). Across all SD3 subscales, fewer participants than anticipated endorsed *neither agree nor disagree*, which suggested that response options could be collapsed to exclude the neutral response category.

Participants were less likely to endorse the *agree* or *strongly agree* response options when they exhibited lower levels of the latent construct and vice versa indicating that

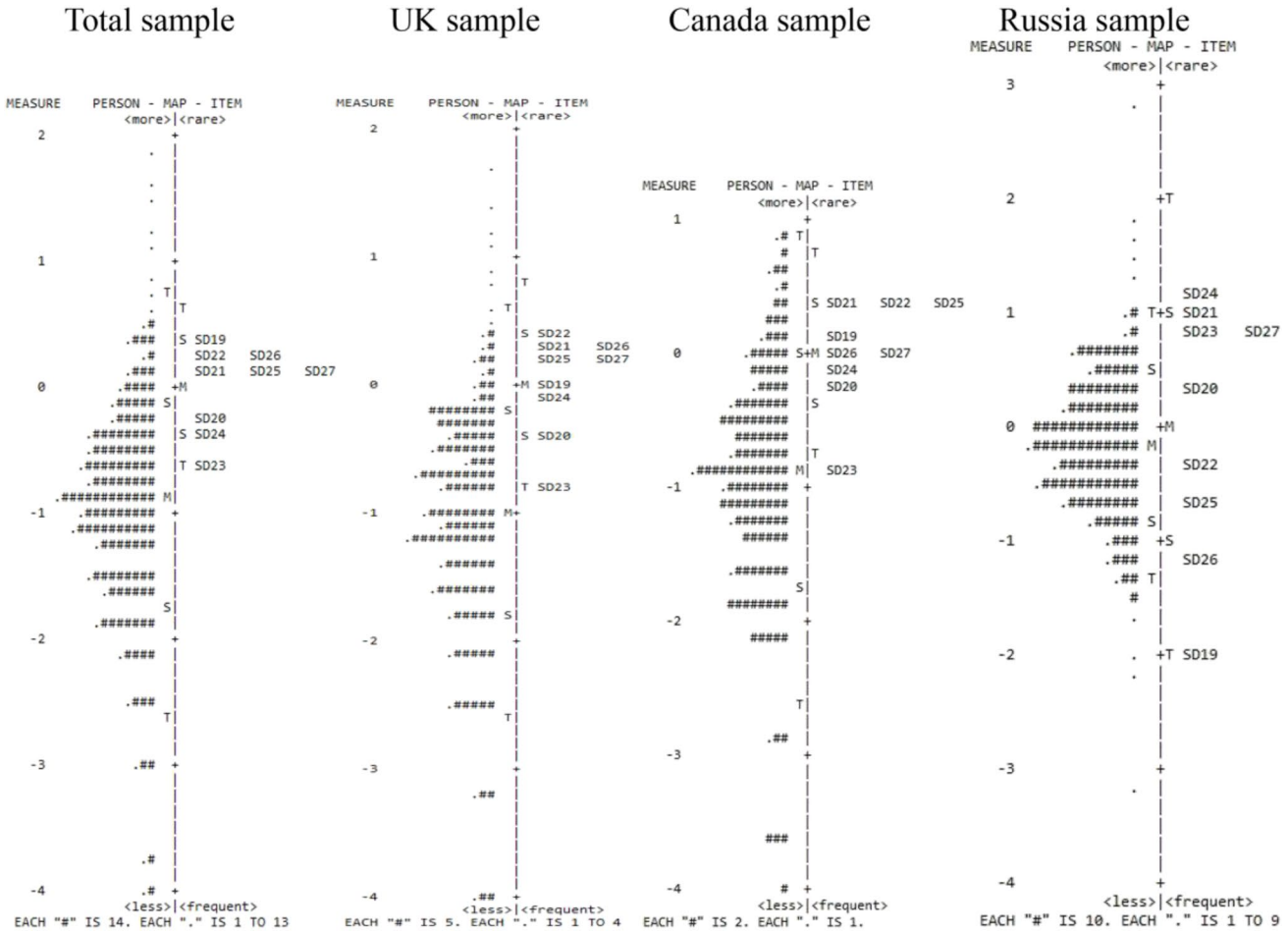


Fig. 4 Person-Item Maps of Psychopathy Across Nations. The participants are on the left of each dashed line and more able participants are located at the top of each map. Items are located on the right of

each dashed line and more difficult items are also located at the top of each map. Note. M=Mean persons' ability or mean items' difficulty; S=one standard deviation; T=two standard deviations

item difficulty values were adequate. Some items were more readily endorsed. These included, for instance, Machiavellianism Item 1 (“It’s not wise to tell your secrets”) and narcissism Item 5 (“I like to get acquainted with important people”). These items did not necessarily reflect explicit features of maladaptive traits. Illustratively, reluctance to disclose personal information may occur in the absence of the tendency to manipulate and exploit others. Hence, this item may assess related but distinct constructs such as self-concealment (Larson and Chastain, 1990). Consistent with past IRT findings (Dinić et al., 2018; Persson et al., 2019), psychopathy items were more difficult to endorse.

of items and/or a sample with a broader range of abilities (Linacre, 2012).

Across countries, some subscales exhibited low person reliability. This suggested that they were unable to effectively discriminate between individuals with varying levels of each latent trait. This may have been a product of the samples employed. Explicitly, the present study recruited community samples that were less likely (vs. clinical or forensic groups) to exhibit maladaptive characteristics. To address this, future studies should assess a larger subset

Lastly, we assessed item equivalence across countries, and found that narcissism and psychopathy items exhibited significant DIF when comparing the Russian sample with the UK and Canada (no notable DIF existed between the UK and Canada). For example, psychopathy items 3 (“Payback needs to be quick and nasty”) and 6 (“People who mess with me always regret it”) were more difficult to endorse for participants in Russia than in the remaining countries. Participants in Russia also found narcissism Item 9 (“I insist on getting the respect I deserve”) more challenging to endorse, and Item 8 (“I am an average person”) easier to endorse. The reasons for DIF are difficult to disentangle; however, this may have been due to translation issues and/or differences in terms of what is considered culturally appropriate across countries.

For example, although narcissism Item 9 seemed to reflect similar phrasing after translation (“I demand to be treated with respect which I surely deserve”), it is plausible

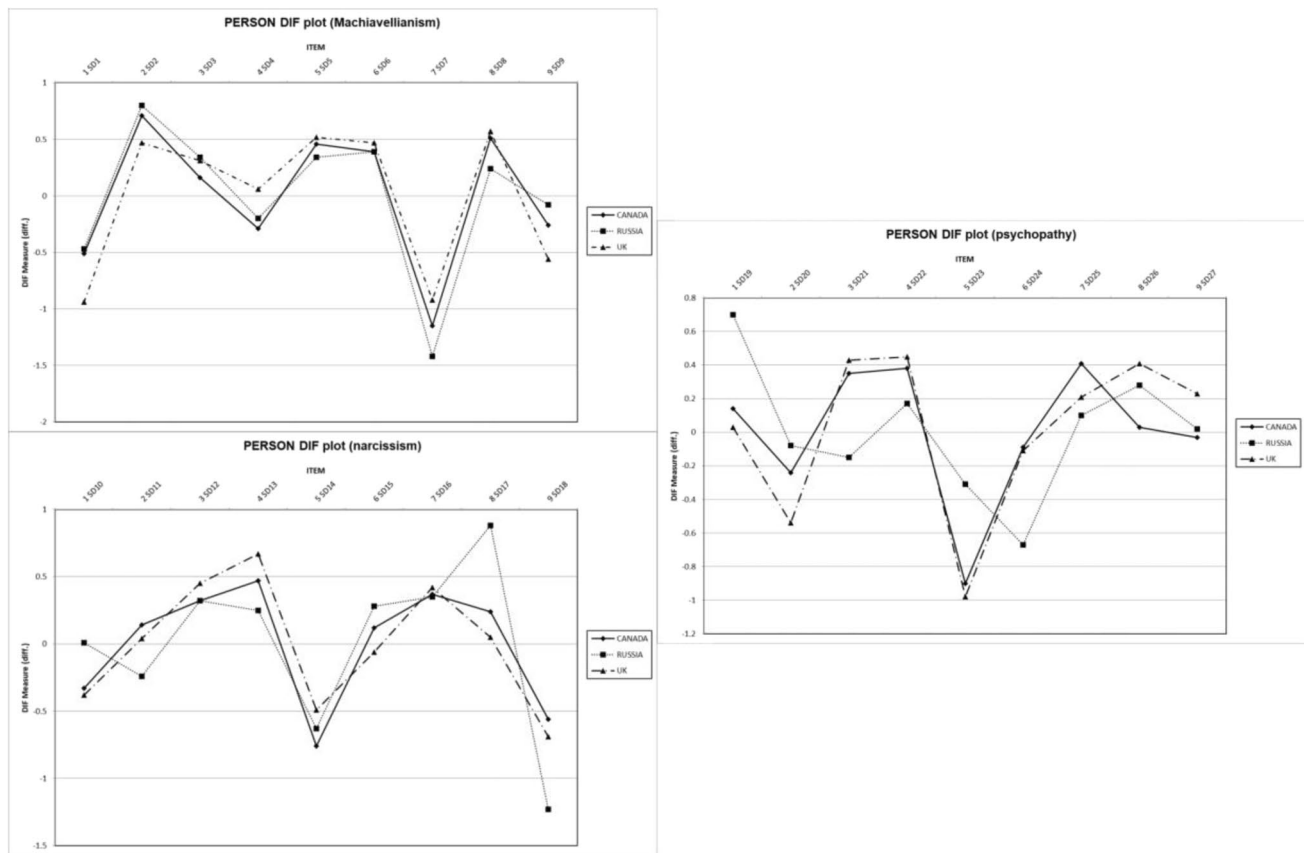


Fig. 5 DIF Plots for SD3 Subscales in Relation to Country

that focusing on being respected is less relevant to one's narcissistic characteristics in Russia, as individuals are generally less individualistic and more interested in preserving collectivistic values (Hofstede, 2001). Other items, conversely, may have reflected discrepancies in meaning. For instance, psychopathy Item 3 translated as “*Accounts must be settled quickly and mercilessly*”. The phrase ‘accounts’ is different in meaning than ‘payback’, which may have impacted responses. The DIF results generally indicate that SD3 items should be revisited to be culturally sound.

Overall, however, Rasch model results were consistent with past IRT SD3 findings. They indicated that each construct was adequately represented by scale items (Dinić et al., 2018; Somma et al., 2019). This outcome contradicted Persson et al. (2019), who proposed that SD3 constructs should not be treated individually because items related to Machiavellianism and psychopathy, as well as narcissism and psychopathy, reflect a blend of Dark Triad constructs. Though this paper found that the SD3 could serve as a valid assessment tool for capturing Dark Triad traits across countries, the observation of insufficient person reliability and DIF indicate that the improvement of SD3 items would subsequently improve measurement precision.

The theoretical and practical implications of this work are far-reaching. Self-report measures of personality are often used to predict important behavioural outcomes, such as counterproductive work behaviours (Forsyth et al., 2012), and relationship aggression (Plouffe et al., 2020). To ensure accurate prediction of these and other relevant behaviours, it is imperative that measures of the Dark Triad are valid and reliable. Moreover, it is conceptually important that the SD3 is available across multiple countries and languages as cross culture comparisons will further deepen understanding of Dark Triad traits.

The findings of this paper broadly aligned with the factor structure proposed by McLarnon and Tarraf (2017), and the view of Jones and Paulhus (2014) that three largely distinct scales exist. Explicitly, the results specify that the SD3 most appropriately captures a Dark Triad, rather than a Dark Dyad or Core. However, some item properties were questionable, and DIF was present across countries. Thus, two important questions remain: Is the SD3 a valid and reliable tool to measure the Dark Triad traits? Are the constructs measured using the SD3 redundant or reasonably distinct? This paper recommends that the measurement properties of the SD3 are not only re-examined in future studies, but also revised to ensure that items possess unequivocally sound

psychometric properties across samples, including effective differentiation between constructs. Items that should be revised include narcissism items 8 and 9, and psychopathy items 1, 3, 5, and 6.

Theoretical and practical contributions

Conclusively, this study supported a three-factor SD3 model. Although the bifactorial model demonstrated greater data-fit, scrutiny revealed that the general factor was insubstantial. This indicated that the Dark Triad within the SD3 is best conceived as three associated constructs. This interpretation is theoretically sound since it aligns with predominant conceptualisations of the Dark Triad.

Use of ESEM resolved measurement issues, which have previously obfuscated clarification of SD3 latent composition. Particularly, bifactor solutions obtain greater data-fit and evidence stronger inter-factor correlations because of the limitations of CFA, which restricts item loading to target factors. This convention explains why researchers have found support for two-factor models. ESEM, due to tolerance of cross-loading, provides a more credible assessment of SD3 structure, supporting a (primarily) multidimensional solution with non-redundant general factor variance.

Thus, future research should use ESEM when modelling SD3 responses. This is necessary since SD3 items are not pure measures of targeted factors and possess variance affiliated with adjacent constructs (Morin et al., 2016). It is important for subsequent work to account for this when using the SD3. Moreover, the SD3 should be examined/employed as a three-construct measure, as opposed to uni- or bidimensional. Relatedly, the Russian version requires refinement and/or further research to disentangle sources of DIF in comparison with the original, English-speaking version. This will facilitate precise use of the SD3 within Russian-based studies when examining Dark Triad relations with other constructs.

Psychometric validation of the SD3 across countries has important implications for educators and policymakers. For educators, SD3 dimensions can inform development of interventions that address challenging behaviours. Specifically, inform strategies that moderate destructive student responses (lack of compassion, deceit, manipulation, etc.). Additionally, educationalists can incorporate content within the curriculum that promotes pro-social behaviours and enhances emotional intelligence/awareness. These measures will reduce the potentially detrimental impact of Dark Triad traits in education settings and nurture positive student behaviour. A concomitant benefit to educators is that familiarity with Dark Triad personality traits will empower them to foster an inclusive and supportive learning environment.

This is important because though Dark Triad traits can prove dysfunction (Kaufman et al., 2019), they are also linked to personal success and advancement (Furnham et al., 2013). Hence, educators need to carefully manage rather than eradicate Dark Triad traits, ensuring that they operate in constructive and adaptive ways. Unchallenged, Dark Triad traits can produce attitudes and behaviours that undermine social relationships and achievement in educational settings. Illustratively, Dark Triad traits correlate negatively with intellectual humility, positively with academic cheating (e.g., Williams et al., 2010), and narcissism and psychopathy are allied to academic entitlement (Turnipseed & Cohen, 2015).

Beyond educational contexts, delineation of Dark Triad traits helps policymakers develop and implement guidelines that promote prosocial behaviour. For instance, in the workplace training can help managers focus on ethical leadership and manage challenging attitudes and conduct (Diller et al., 2021). Awareness of negative traits can thus foster healthier real-world environments. From this perspective, policymakers can produce procedures that mitigate the possible negative effects of Dark Triad traits and advance ethical and empathetic social responses (Furtner et al., 2017). Initiatives are best employed at a structural level since researchers have found individual interventions less successful (Diller et al., 2021). At a practical level, this suggests that policymakers should focus on the reinforcement of ethical values within senior and leadership positions. For instance, within organisations, a culture with core values of mutual respect and cooperation accompanied by transparent rules and career paths can reduce the undesirable consequences of dark dispositions (Crawshaw, 2007).

Readers should view theoretical and practical contributions cautiously due to issues, which potentially affect result generalisability. Particularly, the use of non-random (i.e., convenience) sampling. This method may inadvertently exclude respondents (e.g., forensic and clinical populations; mentioned above) who are important to the study of the Dark Triad. Moreover, classical test statistics (as used for factorial analysis) are sample dependent and estimations change as a function of context (Henson, 1999). Thus, estimations are not necessarily generalisable to different samples. Use of Rasch to produce fit and examine DIF helps to remedy sample variation concerns. This is because Rasch item difficulty statistics are, theoretically, ‘person-free’ (i.e., do not alter if different participants are included) (Henard, 2000).

Limitations and future directions

Though this work has important implications, it is not without its limitations. First, participants were recruited using social networks, word of mouth, and a research participant

pool. It is possible that our results may have differed if participants represented a larger range of trait levels. To assess the psychometric properties of the SD3 more comprehensively, future studies should recruit participants using more diverse methods, such as through clinical groups. Furthermore, the UK and Russian samples included more females than males. Females typically score lower on the SD3 (Jones & Paulhus, 2014), which may be a result of ‘actual’ gender difference or a measurement artefact. It would be important for future research to examine gender in the context of Rasch modelling to clarify if gender DIF exists. Relatedly, the presence of comparably younger participants within the Canada dataset vs. the UK and Russia is a limitation because age differences impact the expression of Dark Triad traits. Precisely, scores peak in early adulthood (18–30 years), and weaken over time (Jonason et al., 2012). Therefore, the Canadian sample may have responded more strongly to the SD3, although DIF analyses revealed no major/significant differences existed in comparison with the UK.

Next, because data relied on self-reporting, there is a potential risk of socially desirable responding, such that participants consciously present themselves more favorably. In future studies, we recommend controlling for social desirability when examining associations between the SD3 and relevant correlates. This can include methods such as implementing attention checks (e.g., as with the HEXACO-PI-R; Ashton & Lee, 2008) and/or including social desirability scales. Indeed, to improve validity within psychometric-based research, social desirability scales can identify, limit, and correct for aberrant responding (van de Mortel, 2008).

Lastly, a limitation includes the lack of related measures (such as the DTDD) to examine concurrent, convergent and discriminant validity of the SD3 across different samples. However, the project provides valuable information concerning the performance of the SD3 at a structural and item level.

Conclusion

Comparison of SD3 item performance across countries is under researched. Hence, researchers have failed to reach consensus about the latent composition of the SD3. This study observed SD3 response deviations within the Russia sample. Additionally, subscales demonstrated low person reliability, indicating insensitivity when distinguishing between high and low scorers. To enhance reliability researchers should consider development of additional items.

Analysis of factor structure using ESEM provided support for a three-factor SD3 conceptualization. Rasch analysis supported subscale unidimensionality. These outcomes

suggest the inherent CFA analytical limitations have conflated previously reported factorial solutions. In this context, ESEM provided a more pragmatic assessment of subscale structure, which future research should utilize when modeling SD3 responses. This approach has the potential to facilitate greater conceptual consensus. To conclude, given the inconsistent SD3 factor structures across studies and DIF across countries, we recommend that researchers consider ways to improve the measure. This may involve rewriting specific items or including additional items designed to measure the Dark Triad traits and maximize differentiation between Machiavellianism and psychopathy.

Authors’ contribution All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Andrew Denovan and Rachel Plouffe. The first draft of the manuscript was written by Andrew Denovan, Rachel Plouffe, and Neil Dagnall, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding No funding bodies to report.

Data availability The study data are available from figshare: <https://doi.org/10.6084/m9.figshare.24018099>.

Declarations

Ethical approval Ethical review boards based in Western University, Canada, and Manchester Metropolitan University, UK, approved the study. The procedures used adhere to the tenets of the Declaration of Helsinki.

Consent Informed consent was obtained from all individual participants included in the research.

Competing interests Rachel Plouffe is an Associate Editor of Current Psychology. The remaining authors have no conflicts of interest to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Amin, L., Rosenbaum, P., Barr, R., Sung, L., Klaassen, R. J., Dix, D. B., & Klassen, A. (2012). Rasch analysis of the PedsQL: An increased understanding of the properties of a rating scale.

- Journal of Clinical Epidemiology*, 65(10), 1117–1123. <https://doi.org/10.1016/j.jclinepi.2012.04.014>
- Ashton, M. C., & Lee, K. (2008). The HEXACO model of personality structure and the importance of the H factor. *Social and Personality Psychology Compass*, 2(5), 1952–1962. <https://doi.org/10.1111/j.1751-9004.2008.00134.x>
- Bandalos, D. L. (2014). Relative performance of categorical diagonally weighted least squares and robust maximum likelihood estimation. *Structural Equation Modeling*, 21(1), 102–116. <https://doi.org/10.1080/10705511.2014.859510>
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.
- Christie, R., & Geis, F. L. (1970). *Studies in Machiavellianism*. Academic Press.
- Cleckley, H. (1976). *The mask of sanity* (5th ed.). Mosby.
- Crawshaw, L. (2007). *Taming the abrasive manager: How to end unnecessary roughness in the workplace*. Wiley.
- Denovan, A., Dagnall, N., Artamonova, E., & Papageorgiou, K. A. (2021). Dark Triad traits, learning styles, and symptoms of depression: Assessing the contribution of mental toughness longitudinally. *Learning and Individual Differences*, 91, 102053. <https://doi.org/10.1016/j.lindif.2021.102053>
- Denovan, A., Dagnall, N., Drinkwater, K., & Escolà-Gascón, Á. (2023). Evaluating the psychometric properties of the chronic time pressure inventory using Rasch analysis. *PeerJ*, 11, e15218. <https://doi.org/10.7717/peerj.15218>
- Diller, S. J., Czibor, A., Szabó, Z. P., Restás, P., Jonas, E., & Frey, D. (2021). The positive connection between dark triad traits and leadership levels in self-and other-ratings. *Leadership Education Personality: An Interdisciplinary Journal*, 1–15. <https://doi.org/10.1365/s42681-021-00025-6>
- 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>
- Doucercain, M. M., Ryder, A. G., & Amiot, C. E. (2021). What are friends for in Russia versus Canada? An approach for documenting cross-cultural differences. *Cross-Cultural Research*, 55(4), 382–409. <https://doi.org/10.1177/10693971211024599>
- Egan, V., Chan, S., & Shorter, G. W. (2014). The Dark Triad, happiness and subjective well-being. *Personality and Individual Differences*, 67, 17–22. <https://doi.org/10.1016/j.paid.2014.01.004>
- Egorova, M. S., Sitnikova, M. A., & Parshikova, O. V. (2015). Adaptation of the short Triad Questionnaire. *Psychological Research: Electronic Scientific Journal*, 8, 1–1.
- Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock, & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269–314). Information Age.
- Forbes, D., Lockwood, E., Creamer, M., Bryant, R. A., McFarlane, A. C., Silove, D., Nickerson, A., & O'Donnell, M. (2015). Latent structure of the proposed ICD-11 post-traumatic stress disorder symptoms: Implications for the diagnostic algorithm. *The British Journal of Psychiatry*, 206(3), 245–251. <https://doi.org/10.1192/bjp.bp.114.150078>
- Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS estimation. *Structural Equation Modeling*, 16(4), 625–641. <https://doi.org/10.1080/10705510903203573>
- 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(3), 557–579. <https://doi.org/10.1037/a0025679>
- Furnham, A., Richards, S. C., & Paulhus, D. L. (2013). The Dark Triad of personality: A 10 year review. *Social and Personality Psychology Compass*, 7(3), 199–216. <https://doi.org/10.1016/j.paid.2014.02.001>
- Furtner, M. R., Maran, T., & Rauthmann, J. F. (2017). Dark leadership: The role of leaders' dark triad personality traits. *Leader Development Deconstructed*, 75–99. https://doi.org/10.1007/978-3-319-64740-1_4
- Glenn, A. L., & Sellbom, M. (2015). Theoretical and empirical concerns regarding the dark triad as a construct. *Journal of Personality Disorders*, 29(3), 360–377. https://doi.org/10.1521/pedi_2014_28_162
- Han, H. (2022). The effectiveness of weighted least squares means and variance adjusted based fit indices in assessing local dependence of the Rasch model: Comparison with principal component analysis of residuals. *PLoS ONE*, 17(9), e0271992. <https://doi.org/10.1371/journal.pone.0271992>
- Henard, D. H. (2000). Item response theory. In L. Grimm, & P. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 67–97). American Psychological Association.
- Henson, R. (1999, January). *Understanding the one-parameter Rasch model of item response theory*. Paper presented at the annual meeting of the Southwest Educational Research Association, San Antonio, TX. (ERIC Document Reproduction Service No. ED 428 078).
- Hodson, G., Book, A., Visser, B. A., Volk, A. A., Ashton, M. C., & Lee, K. (2018). Is the dark triad common factor distinct from low honesty-humility? *Journal of Research in Personality*, 73, 123–129. <https://doi.org/10.1016/j.jrp.2017.11.012>
- Hofstede, G. (2001). Culture's recent consequences: Using dimension scores in theory and research. *International Journal of Cross-Cultural Management*, 1(1), 11–17. <https://doi.org/10.1177/147059580111002>
- Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment*, 22(2), 420–432. <https://doi.org/10.1037/a0019265>
- Jonason, P. K., Slomski, S., & Partyka, J. (2012). The Dark Triad at work: How toxic employees get their way. *Personality and Individual Differences*, 52(3), 449–453. <https://doi.org/10.1016/j.paid.2011.11.008>
- Jonason, P. K., Foster, J., Oshio, A., Sitnikova, M., Birkas, B., & Gouveia, V. (2017). Self-construals and the Dark Triad traits in six countries. *Personality and Individual Differences*, 113, 120–124. <https://doi.org/10.1016/j.paid.2017.02.053>
- Jonason, P. K., Sitnikova, M., & Oshio, A. (2018). The Dark Triad traits and views of time in three countries. *Personality and Individual Differences*, 120, 107–111. <https://doi.org/10.1016/j.paid.2017.08.036>
- Jonason, P. K., Žemojtel-Piotrowska, M., Piotrowski, J., Sedikides, C., Campbell, W. K., Gebauer, J. E., Maltby, J., Adamovic, M., Adams, B. G., Kadiyono, A. L., Atitsogbe, K. A., Bundhoo, H. Y., Băltătescu, S., Bilić, S., Brulin, J. G., Chobthamkit, P., Del Carmen Dominguez, A., Dragova-Koleva, S., El-Astal, S., & Yahiaev, I. (2020). Country-level correlates of the dark triad traits in 49 countries. *Journal of Personality*, 88(6), 1252–1267. <https://doi.org/10.1111/jopy.12569>
- Jones, D. N., & Paulhus, D. L. (2014). Introducing the short dark triad (SD3) a brief measure of dark personality traits. *Assessment*, 21, 28–41. <https://doi.org/10.1177/1073191113514105>
- Kaufman, S. B., Yaden, D. B., Hyde, E., & Tsukayama, E. (2019). The light vs. dark triad of personality: Contrasting two very different profiles of human nature. *Frontiers in Psychology*, 10, 467. <https://doi.org/10.3389/fpsyg.2019.00467>

- Kowalski, C. M., Di Pierro, R., Plouffe, R. A., Rogoza, R., & Saklofske, D. H. (2020). Enthusiastic acts of evil: The Assessment of sadistic personality in Polish and Italian populations. *Journal of Personality Assessment, 102*(6), 770–780. <https://doi.org/10.1080/00223891.2019.1673760>
- 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*, 103455. <https://doi.org/10.1016/j.actpsy.2021.103455>
- Larson, D. G., & Chastain, R. L. (1990). Self-concealment: Conceptualization, measurement, and health implications. *Journal of Social and Clinical Psychology, 9*(4), 439–455. <https://doi.org/10.1521/jscp.1990.9.4.439>
- Linacre, J. M. (2012). *Dimensionality: Contrasts & variances*. Retrieved from <http://www.winsteps.com/winman/index.htm?principalcomponents.htm>
- Ludlow, L. H., Enterline, S. E., & Cochran-Smith, M. (2008). Learning to teach for social justice-beliefs scale: An application of Rasch measurement principles. *Measurement and Evaluation in Counseling and Development, 40*(4), 194–214. <https://doi.org/10.1080/07481756.2008.11909815>
- Maples, J. L., Lamkin, J., & Miller, J. D. (2014). A test of two brief measures of the dark triad: The dirty dozen and short dark triad. *Psychological Assessment, 26*(1), 326–331. <https://doi.org/10.1037/a0035084>
- Marsh, H. W., Vallerand, R. J., Lafrenière, M. A. K., Parker, P., Morin, A. J., Carbonneau, N., Jowett, S., Bureau, J. S., Fernet, C., Frédéric, G., Salah Abduljabbar, A., & Paquet, Y. (2013). Passion: Does one scale fit all? Construct validity of two-factor passion scale and psychometric invariance over different activities and languages. *Psychological Assessment, 25*(3), 796–809. <https://doi.org/10.1037/a0032573>
- McHoskey, J. W., Worzel, W., & Szyarto, C. (1998). Machiavellianism and psychopathy. *Journal of Personality and Social Psychology, 74*(1), 192–210. <https://doi.org/10.1037/0022-3514.74.1.192>
- McLarnon, M. J. (2022). Into the heart of darkness: A person-centered exploration of the Dark Triad. *Personality and Individual Differences, 186*, 111354. <https://doi.org/10.1016/j.paid.2021.111354>
- McLarnon, M. J., & Tarraf, R. C. (2017). The Dark Triad: Specific or general sources of variance? A bifactor exploratory structural equation modeling approach. *Personality and Individual Differences, 112*, 67–73. <https://doi.org/10.1016/j.paid.2017.02.049>
- Morin, A. J., Marsh, H., & Nagengast, B. (2013). Exploratory structural equation modeling. In G. R. Hancock, & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 395–436). Information Age Publishing.
- Morin, A. J. S., Arens, A. K., Tran, A., & Caci, H. (2016). Exploring sources of construct-relevant multidimensionality in psychiatric measurement: A tutorial and illustration using the composite scale of morningness. *International Journal of Methods in Psychiatric Research, 25*, 277–288. <https://doi.org/10.1002/mpr.1485>
- Muthén, L. K., & Muthén, B. O. (2018). *Mplus user's guide* (8th ed.). Muthén & Muthén.
- Papageorgiou, K. A., Denovan, A., Dagnall, N., Hill-Artamonova, E., Gianniu, F. M., Papageorgiou, S., Plouffe, R. A., Kowalski, C. M., Saklofske, D. H., Kyriazos, T., Stalikas, A., & Costantini, G. (2023). Grandiose narcissism indirectly associates with lower psychopathology across five countries. *Journal of Psychiatric Research, 167*, 78–85. <https://doi.org/10.1016/j.jpsychires.2023.10.003>
- Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, machiavellianism, and psychopathy. *Journal of Research in Personality, 36*(6), 556–563. [https://doi.org/10.1016/S0092-6566\(02\)00505-6](https://doi.org/10.1016/S0092-6566(02)00505-6)
- Paulhus, D. L., Neumann, C. S., Hare, R. D., Williams, K. M., & Hemphill, J. F. (2016). *Self-Report Psychopathy Scale—Technical manual*. Multi-Health Systems.
- Pechorro, P., Caramelo, V., Oliveira, J. P., Nunes, C., Curtis, S. R., & Jones, D. N. (2019). The short Dark Triad (SD3): Adaptation and psychometrics among at-risk male and female youths. *Deviant Behavior, 40*(3), 273–286. <https://doi.org/10.1080/01639625.2017.1421120>
- Persson, B. N., Kajonius, P. J., & Garcia, D. (2017). Testing construct independence in the short dark triad using item response theory. *Personality and Individual Differences, 117*, 74–80. <https://doi.org/10.1016/j.paid.2017.05.025>
- Persson, B. N., Kajonius, P. J., & Garcia, D. (2019). Revisiting the structure of the short Dark Triad. *Assessment, 26*(1), 3–16. <https://doi.org/10.1177/1073191117701192>
- Pineda, D., Sandín, B., & Muris, P. (2020). Psychometrics properties of the Spanish version of two dark triad scales: The dirty dozen and the short Dark Triad. *Current Psychology, 39*(5), 1873–1881. <https://doi.org/10.1007/s12144-018-9888-5>
- Plouffe, R. A., Wilson, C. A., & Saklofske, D. H. (2020). The role of dark personality traits in intimate partner violence: A multi-study investigation. *Current Psychology, 41*, 3481–3500. <https://doi.org/10.1007/s12144-020-00871-5>
- Plouffe, R. A., Kowalski, C. M., Papageorgiou, K. A., Dinić, B. M., Artamonova, E., Dagnall, N., Denovan, A., Gianniu, F. M., Kyriazos, T., Saklofske, D. H., & Stalikas, A. (2023). The revised Assessment of sadistic personality (ASP-8): Evidence for validity across four countries. *Journal of Personality Assessment, 105*(2), 149–162. <https://doi.org/10.1080/00223891.2022.2055476>
- Raskin, R. N., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports, 45*(2), 590–590. <https://doi.org/10.2466/pr0.1979.45.2.590>
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment, 98*(3), 223–237. <https://doi.org/10.1080/00223891.2015.1089249>
- Rogoza, R., & Cieciuch, J. (2019). Structural investigation of the short Dark Triad questionnaire in Polish population. *Current Psychology, 38*(3), 756–763. <https://doi.org/10.1007/s12144-017-9653-1>
- Shirokova, G., Selivanovskikh, L., Bordunos, A., & Morris, M. H. (2022). Examining the Role of the Dark Side of CEO Personality in Decision-Making Approaches Within Russia. In *Academy of Management Proceedings* (Vol. 2022, No. 1, p. 12748). Briarcliff Manor, NY 10510: Academy of Management. <https://doi.org/10.5465/AMBPP.2022.348>
- Siddiqi, N., Shahnawaz, M., & Nasir, S. (2020). Reexamining construct validity of the short Dark Triad (SD3) scale. *Current Issues in Personality Psychology, 8*(1), 18–30. <https://doi.org/10.5114/cipp.2020.94055>
- Somma, A., Paulhus, D. L., Borroni, S., & Fossati, A. (2019). Evaluating the psychometric properties of the short Dark Triad (SD3) in Italian adults and adolescents. *European Journal of Psychological Assessment, 36*(1), 185–195. <https://doi.org/10.1027/1015-5759/a000499>
- Southard, A. C., & Zeigler-Hill, V. (2016). Measuring the dark side of personality. In U. Kumar (Ed.), *The Wiley handbook of personality assessment* (pp. 119–133). Wiley. <https://doi.org/10.1002/9781119173489.ch9>
- Turnipseed, D. L., & Cohen, S. R. (2015). Academic entitlement and socially aversive personalities: Does the Dark Triad predict academic entitlement? *Personality and Individual Differences, 82*, 72–75. <https://doi.org/10.1016/j.paid.2015.03.003>
- Valentova, J. V., Junior, F. P. M., Štěrbová, Z., Varella, M. A. C., & Fisher, M. L. (2020). The association between Dark Triad traits and sociosexuality with mating and parenting efforts: A

- cross-cultural study. *Personality and Individual Differences*, 154, 109613. <https://doi.org/10.1016/j.paid.2019.109613>
- Van de Mortel, T. F. (2008). Faking it: Social desirability response bias in self-report research. *The Australian Journal of Advanced Nursing*, 25(4), 40–48. <https://doi.org/10.3316/informit.210155003844269>
- Wetzel, E., Lang, F. J., Back, M. D., Vecchione, M., Rogoza, R., & Roberts, B. W. (2020). Measurement invariance of three narcissism questionnaires across the United States, the United Kingdom, and Germany. *Assessment*, 28(1), 29–43.
- Williams, K. M., Nathanson, C., & Paulhus, D. L. (2010). Identifying and profiling scholastic cheaters: Their personality, cognitive ability, and motivation. *Journal of Experimental Psychology: Applied*, 16(3), 293–307. <https://doi.org/10.1037/a0020773>
- Wright, B., & Linacre, J. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8, 370–371.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.