RESEARCH ARTICLE

Personality and Mental Health

Assessing the dark personality traits with observer reports: A meta-analysis of inter-rater agreement on the Dark Triad and Dark Tetrad traits

Pilar Rico-Bordera[1] | David Pineda[1] | Manuel Galán[1,2] | José A. Piqueras[1]

[1]Forensic Psychology Unit of the Centre for Applied Psychology, Miguel Hernández University of Elche, Avenida de la Universidad, Alicante, Spain

[2]Department of Psychology, Faculty of Medicine, Catholic University of Murcia, Guadalupe de Maciascoque, Murcia, Spain

Correspondence
David Pineda, Forensic Psychology Unit of the Centre for Applied Psychology, Miguel Hernández University of Elche, Avenida de la Universidad, s/n, 03202 Alicante, Spain.
Email: dpineda@umh.es

Abstract

Assessing dark personality traits with self-reports may be hampered by biases such as socially desirable responding. To address this, observer reports could be an effective complement or alternative to self-reports. This meta-analysis aimed to investigate the correlations of observer reports of the Dark Triad and Dark Tetrad traits with these traits' self-reports. A comprehensive search in PubMed, Web of Science, Scopus and PsycINFO yielded 24 studies that met the inclusion criteria (with 32 effect sizes) and included the assessment of at least one of the dark traits. Results showed positive self-observer correlations of medium-high magnitude for psychopathy ($r = 0.47$) and of medium magnitude for narcissism and Machiavellianism (respectively, $r = 0.44$ and 0.41). In addition, for Machiavellianism and psychopathy, it was shown that the better acquainted the observer and the observed, the higher the self-observer correlations. For sadism, no results could be obtained because only one study included the assessment of this trait. This is the first meta-analysis of inter-rater agreement in assessing the Dark Tetrad. The findings suggest that observer reports could be a supplement or even be alternatives to self-reports when assessing the Dark Tetrad traits.

INTRODUCTION

Several studies have examined the value of observer reports for assessing the Big Five and HEXACO (model of the six major traits) personality traits. They have shown that observer reports correlate at least moderately with self-reports and provide comparable—and sometimes even higher—levels of predictive validity for important outcomes (Connelly & Ones, 2010; Connelly et al., 2007; Kim et al., 2019). This suggests that observer reports can be useful supplements or even alternatives to self-reports (Abernethy, 2015; Connelly & Ones, 2010; Connelly et al., 2007; Kim et al., 2019; Vazire & Carlson, 2011). For example, since about 40% of trait score variance in self-reports is specific to this method (McCrae, 2018), research findings based on self-reports alone could be biased and should be validated using alternative methodologies such as observer reports (Abernethy, 2015; Hofstee, 1994; McDonald, 2008; Vazire & Carlson, 2011). Moreover, combining self-reports and observer reports can often yield much stronger effects sizes such as heritability estimates (Kandler et al., 2010; Riemann et al., 1997) or personality traits' associations with other variables (e.g., Muris et al., 2013).

These empirical observations are supplemented by theoretical models on the accuracy of person perception. For example, the realistic accuracy model (RAM; Funder, 1995) articulates factors contributing to the

accuracy of personality judgements by observers. These factors include relevance, availability, detection and utilization of behavioural cues, based on the idea that accuracy in personality judgement is a joint product of the target's attributes and behaviour, as well as the judge's observations and insight. Relatedly, the self-other knowledge asymmetry model (SOKA; Vazire, 2010) postulates that the self-ratings may be more accurate than observer ratings for assessing comparatively less visible and less evaluative traits, whereas observer reports can be more accurate for assessing highly observable and evaluative traits.

The dark personality traits

The Dark Tetrad is a set of four malevolent traits that, although distinct, are closely associated with each other and all highly evaluative (Chabrol et al., 2009; Jones & Paulhus, 2011; Paulhus & Williams, 2002). Narcissism is characterized by egocentric self-admiration and feelings of grandiosity, which produce feelings of superiority and the need for ego reinforcement (Jones & Paulhus, 2014; Paulhus & Williams, 2002; Raskin & Hall, 1979). Machiavellianism is characterized by callous affect and manipulation, which is conducive to a strategic orientation focused on self-interest and personal gain (Christie & Geis, 1970; Jones & Paulhus, 2009, 2014; Paulhus & Williams, 2002). Psychopathy is characterized by antisocial behaviour, impulsivity, sensation-seeking, low empathy and lack of remorse (Hare, 1970; Jones & Paulhus, 2014; Paulhus & Williams, 2002). These three traits are also known as the Dark Triad (Paulhus & Williams, 2002). The fourth trait, sadism, is represented by the feeling of enjoyment or sense of pleasure in observing or causing harm to others (Chabrol et al., 2009; O'Meara et al., 2011).

Assessing these traits exclusively with self-reports may be particularly problematic because people who score high on these traits may be especially likely to present themselves in a socially desirable way to conceal information about their true trait levels (Andrews & Meyer, 2003; Echeburúa et al., 2011; Spaans et al., 2017). As a result, authors have suggested the use of other methods besides self-reports to assess these dark traits (Cattell & Warburton, 1967; Hernández-López et al., 1999; Lozano-Bleda et al., 2010; Ortner & Proyer, 2015; Rubio et al., 2004; Santacreu & Hernández, 2018). As a result, observer reports could be particularly useful for the assessments of these traits (McDonald, 2008). However, to date, few studies have used observer reports to assess the Dark Triad traits (e.g., Lämmle et al., 2021; Malesza & Kaczmarek, 2020;

Vander Molen et al., 2018). Moreover, much less is known about the degree to which observer reports of the dark traits correlated with self-reports than we know about self-observer agreement on the Big Five and HEXACO traits. Due to the evaluativeness of these traits, one could expect that their self-observer correlations are lower than those for the Big Five and HEXACO traits, which typically average at $r = 0.30$ (Connelly & Ones, 2010). The lower the cross-rater correlations, the stronger the case for assessing the dark traits with multiple methods to improve the assessments' validity (Connelly & Ones, 2010; Connolly et al., 2007; Kim et al., 2019; McDonald, 2008).

The present study

We performed a meta-analysis on the self-observer correlations of the Dark Triad and Dark Tetrad personality traits. We also analysed the possible mediating role of proximity between the observer and the self on self-observer correlations, as several studies have indicated that the self-observer agreement varies with the proximity between the self and the observer: The better acquainted the observer and the observed, the higher the correlation (Connelly & Ones, 2010; Lämmle et al., 2021; Vazire, 2006; Vazire & Mehl, 2008).

METHOD

Search strategy and study selection

A systematic search was conducted in October 2023 in four databases: PubMed, Web of Science, Scopus and PsycINFO. According to the objective and to ensure that the traits assessed in the included papers were those considered dark traits of the Dark Triad or Dark Tetrad, the search terms utilized were Dark Triad, Dark Tetrad, Dark traits, Narcissism, Machiavellianism and Psychopathy, combined to create the following search string: ('Dark Triad' OR 'Dark Tetrad' OR 'Dark traits') OR ('Narcissism' AND 'Machiavellianism' AND 'Psychopathy'). We designated 2002 as the initial year because this is when the set of traits that form the Dark Triad was first presented (Paulhus & Williams, 2002).

The review was pre-registered in PROSPERO, the international prospective register of systematic reviews (registration number: CRD42023393204; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=393204). However, there were some modifications with respect to the preregistration: The preregistration occurred after the search had been performed and the

primary studies had already been examined to some extent; 2021 was recorded as the year of the search, but later, the search was updated in order to broaden the search and obtain more current results, so the actual year was 2023; initially, it was not considered as an objective to analyse the mediating role of proximity between the observer and the self on self-observer correlations; finally, the software used for data analysis was Jamovi (version 2.2.5).

Inclusion and exclusion criteria

For this meta-analysis, only observer reports used to assess Dark Triad or Dark Tetrad traits as initially conceived by Paulhus and Williams (2002) were of interest. Thus, those studies that divided the traits into subfactors (e.g., Machiavellianism in interpersonal tactics, cynical view of human nature and disregard for conventional morality) or that only reported a total Dark Triad or Dark Tetrad dark score (without reporting the score for each trait) were excluded.

However, as an exception, the splitting of the trait narcissism into the vulnerability and grandiosity subfactors was not considered grounds for exclusion because, according to the authors of one of the most commonly used scales to assess the Dark Triad (i.e., the Short Dark Triad, Jones & Paulhus, 2014), the grandiosity subfactor represents the narcissism of the Dark Triad, so it might be a mistake to exclude studies that divided narcissism into vulnerability and grandiosity and not consider the latter. Therefore, unlike with the rest of the dark traits, if a study presented such a division, grandiose narcissism would be of interest to this meta-analysis.

To determine the eligibility of studies, the following specific inclusion criteria were established: (1) studies written in Spanish or English; (2) studies related to the traits of interest (i.e., the Dark Triad or the Dark Tetrad); (3) primary research only, excluding any narrative, systematic, meta-analytic reviews or umbrella reviews; (4) any type of paper, as long as it is complete; (5) studies that divided the Dark Triad and Dark Tetrad into their traits, without presenting a single dark score (e.g., only a total Dark Tetrad); (6) studies that did not split all or some of the traits into subfactors (e.g., they provided the Machiavellianism trait and not its subfactors, i.e., interpersonal tactics, cynical view of human nature and disregard for conventional morality); (7) studies that employed a self-report questionnaire/s to assess at least one of the Dark Triad or Dark Tetrad traits; and (8) studies that used an observer report to assess at least one of the Dark Triad or Dark Tetrad traits.

Screening phases and data extraction

To facilitate the work of eliminating duplicates, all the studies that each of the databases (i.e., PubMed, Scopus, Web of Science and PsycINFO) returned as a result of the search were downloaded into the Zotero reference manager (https://www.zotero.org/).

After eliminating the duplicates ($n = 5991$), the remaining studies were transferred to an Excel sheet to facilitate the next phases of the process, that is, the phases of study selection to eliminate all the works that did not meet the inclusion criteria for meta-analysis. For this purpose, first, the titles and abstracts of each study were read from the Excel sheet, discarding those that did not meet the inclusion criteria. If the reviewers had doubts, the study was not discarded but was considered for the next phase. Secondly, each study not discarded in the first screening was carefully read to consider whether it was definitively included for meeting the established inclusion criteria. After both screening phases, a total of 28 manuscripts were obtained that met the inclusion criteria. However, during the data extraction phase, four of the 28 studies were discarded because, although they did employ observer reports, the study samples were derived from another study. That is, using the same sample, different studies were published, and in each study, different data were reported. Therefore, although this was not an exclusion criterion established a priori, to avoid duplicating the sample and biasing the results, the studies that did not report the values of interest or were published later were discarded (Diller et al., 2021; Holtzman, 2011; Holtzman & Strube, 2013a, 2013b; Hudek-Knezevic et al., 2023; Kardum et al., 2022; Szabó et al., 2023) (Figure 1).

In each of these two phases, two independent reviewers first examined the same studies (10% of all studies) to determine and ensure inter-reviewer agreement. Second, each reviewer examined half of the remaining articles (i.e., 50%) with the collaboration of trained undergraduate students (mentioned in Acknowledgements).

After obtaining the studies that met the inclusion criteria for the meta-analysis, the data of interest (variable coding phase) were extracted in an Excel spreadsheet that would later constitute Table 1. To avoid coding problems and possible discordance between reviewers, a coding manual was prepared detailing the procedure for coding each variable. The data extracted from each study were as follows: (1) authors and year of publication of the study; (2) information on the study sample—participants: country, sample size, age (mean and standard deviation), gender proportion and type of
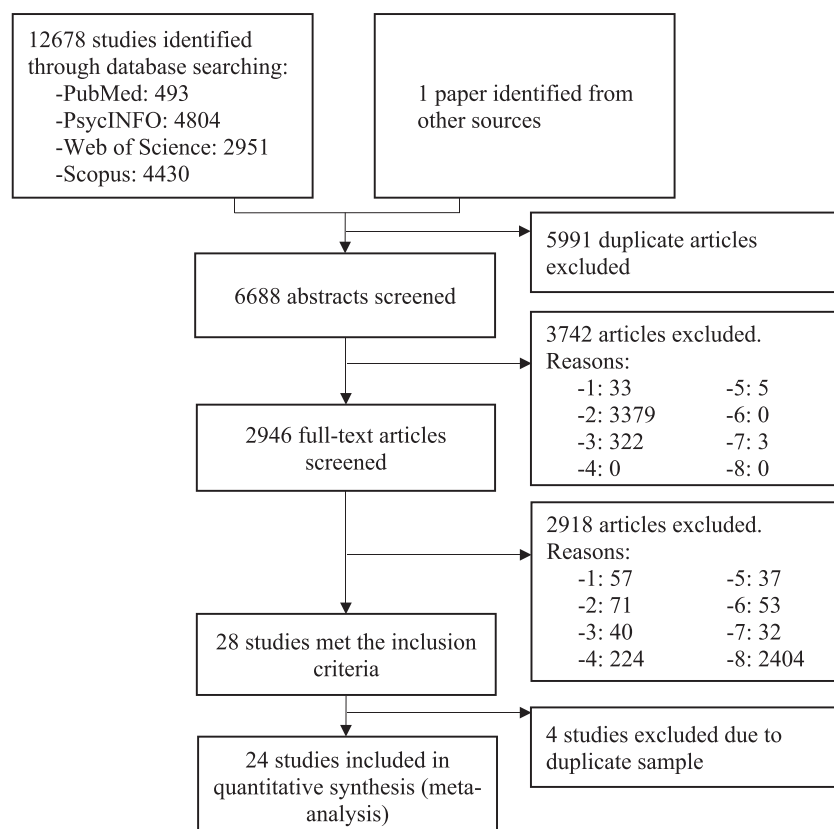
FIGURE 1 Flowchart of the data collection (Preferred Reporting Items for Systematic reviews and Meta-Analyses [PRISMA] flow diagram). *Note.* Exclusion criteria: 1 = language different from English or Spanish; 2 = unrelated to the Dark Triad or Dark Tetrad; 3 = not primary investigation; 4 = full document or relevant information not available; 5 = reports a total Dark Triad or Dark Tetrad score, without dividing it into its traits; 6 = divides all or some of the traits into subfactors without giving a total measure of each trait; 7 = no questionnaire is administered to assess the Dark Triad or Dark Tetrad traits; 8 = does not use an observer report to assess the Dark Triad or Dark Tetrad traits.

sample; (3) information on the questionnaires of interest used: self-report used to measure the Dark Triad or Dark Tetrad traits and its reliability index (i.e., Cronbach's alpha or McDonald's Omega), and observer report used to measure the Dark Triad or Dark Tetrad traits, together with its reliability index (Cronbach's alpha or McDonald's omega) and the informant (i.e., the persons completing the questionnaire); and (4) results obtained: correlation value between the Dark Triad or Dark Tetrad traits measured with the self-report and the observer report.

To ensure correct data extraction in this last phase, two reviewers independently coded the variables of three of the 24 manuscripts (approximately 10% of all studies), and a correlational analysis was performed to obtain an inter-rater reliability value.

Maximum agreement between the two raters was obtained: $r = 1$, $p < 0.001$. Next, one of the previous reviewers coded the variables of the remaining 21 studies, and, in case of doubts with any variable of any study, the rest of the paper's reviewers were informed to make a consensual decision. If any of the data of interest were not reported in the study, the primary reviewer of the study contacted the corresponding author to request the data. All the steps are depicted in Figure 1, which illustrates the PRISMA flow (Moher et al., 2010).

Quality assessment

The methodological quality of the studies was assessed using the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) list (Vandenbroucke et al., 2007; von Elm et al., 2007), composed of 34 checklist criteria of the content that a published study should include. For the cross-sectional studies (those included in this meta-analysis), the list consisted of a total of 32 criteria, scored 1 when the guideline was included in its entirety in the study, 0.50 when it was partially included (incomplete or not very specific) and 0 when it was not included.

These points allowed us to obtain a total quality score for the seven studies. Each study could be classified in four different ways: excellent quality when the score was equal to or greater than 85 (as a result of adding the scores assigned to each item, multiplying them by 100 and dividing by the maximum score that could be obtained, considering that in some cases, 'NA' could be assigned if the item did not apply to the study); good quality when the score was between 70 and 85; fair quality when the score was between 50 and 70; and poor quality when the score was less than 50.

To ensure reliability in quality assessment, again, two independent reviewers analysed three of the 24 studies

TABLE 1 Information about the articles included in the analyses.

Study	Country	Size (N)	Sample age (M; SD)	Gender proportion (female)	Sample type	Self-report (α or ω)	Observer report (α or ω)	Informant	P	N	M	S	STROBE score
Holtzman and Strube (2013a)	United States	151	19.40; 1.22	56%	Students	SRP-III (0.92); NPI-40 (0.85); Mach-IV (0.81)	Ad hoc (0.39); ad hoc (0.78); ad hoc (−0.04)	Friends, acquaintance from college, hometown or high school, current and ex-intimate partner	0.33	0.48	0.26	NI	71.43
Muris et al. (2013)	Netherlands	117	13.90; 0.96	56.41%	Students	DD-Y (0.67); DD-Y (0.66); DD-Y (0.70)	DD-Y (0.71); DD-Y (0.74); DD-Y (0.76)	Parents	0.23	−0.15	0.32	NI	58.93
Jones and Paulhus (2014)	United States and Canada	65	20.10; NI	60%	Mturk	SD3 (0.80); SD3 (0.71); SD3 (0.77)	SD3 (0.86); SD3 (0.67); SD3 (0.62)	Friends, family, romantic partners	0.57	0.34	0.42	NI	62.50
Nealis et al. (2016)	Canada	155	20.70; 3	76.77%	Students	DD (0.78)	DD (0.89)	Friends, family, romantic partners, other relationships	–	0.38	–	–	80.36
Maaß and Ziegler (2017)	Germany	219	37.52; 16.93	64.38%	General population	SD3 (0.70)	NPI-16 (NI)	Trained research assistants	–	0.32	–	–	64.29
Miller et al. (2017)	United States	178	19.30; 2.20	63.92%	Students	SRP-III (0.93); NPI (0.82); Mach-IV (0.81)	SD3 (0.78) + DD (0.82); SD3 (0.63) + DD (0.89); SD3 (0.68) + DD (0.86)	High school and college friends, roommates	0.29	0.17	0.33	NI	51.97
He et al. (2018)	China	260	26.61; 3.96	100%	General population	LSRP (NI); Mach-IV (NI)	LSRP (NI); Mach-IV (NI)	Romantic partners	0.36	–	0.34	–	62.50
	China	260	28.28; 3.84	0%	General population	LSRP (NI); Mach-IV (NI)	LSRP (NI); Mach-IV (NI)	Romantic partners	0.41	–	0.26	–	
Klipfel and Kosson (2018)	United States	62	26.40; 6.70	0%	Prison inmates	NPI (0.85)	IM-N (0.94)	Trained research assistants	–	0.29	–	–	71.43
Vander Molen et al. (2018)	United States	145	20.37; 2.88	82.75%	Students	SD3 + Brief DT (0.80); SD3 + Brief DT (0.76); SD3 + Brief DT (0.80)	Brief DT (0.68); Brief DT (0.82); Brief DT (0.75)	Other people on Facebook	0.09	0.18	0.12	NI	64.29
Weiss et al. (2018)	United States	172	26; 3.40	100%	General population	NEO-FFI ad hoc version (0.73)	NEO-FFI ad hoc version (0.75)	Romantic partners	0.41	–	–	–	62.50

(Continues)

TABLE 1 (Continued)

| Study | Observation samples | | | | | Instruments | | | Results (r) | | | | STROBE score |
	Country	Size (N)	Sample age (M; SD)	Gender proportion (female)	Sample type	Self-report (α or ω)	Observer report (α or ω)	Informant	P	N	M	S	
	United States	172	27.60; 3.90	0%	General population	NEO-FFI ad hoc version (0.70)	NEO-FFI ad hoc version (0.79)	Romantic partners	0.46	–	–	–	
Heinze et al. (2020)	Germany	207	27.68; 9.84	81.48%	General population	NPI (NI)	NPI (NI)	Family, friends, romantic partners, colleague at work, fellow student, roommate, others	–	0.57	–	–	66.07
Malesza and Kaczmarek (2020)	Germany	266	22.90; 1.30	68%	Students	SRP-III (0.79); NPI-17 (0.77); Mach-IV (0.83)	SRP-III (0.98); NPI-17 (0.93); Mach-IV (0.97)	Acquaintance, roommate, significant other	0.46	0.33	0.49	NI	64.29
Kay (2021)	United States	397	19.49; 2.45	72.80%	Students	NPI-40 (0.92)	NPI-16 (0.89)	Family, friends, romantic partners	–	0.39	–	–	78.57
	United States	397	19.49; 2.45	72.80%	Students	SD3 (0.70)	NPI-16 (0.89)	Family, friends, romantic partners	–	0.33	–	–	
Lämmle et al. (2021)	Germany	279	21.44; 1.81	56%	Students	SRP-III (0.73–0.77); NPI (0.77–0.85); Mach-IV (0.43–0.55)	SRP-III (0.76–0.79); NPI (0.75–0.81); Mach-IV (0.44–0.69)	Parents	0.43[a]	0.63[a]	0.26[a]	NI	80.36
	Germany	290	21.44; 1.81	56%	Students	SRP-III (0.73–0.77); NPI (0.77–0.85); Mach-IV (0.43–0.55)	SRP-III (0.80–0.84); NPI (0.84–0.86); Mach-IV (0.45–0.63)	Friends	0.53[a]	0.70[a]	0.28[a]	NI	
Weiss et al. (2021)	United States	408	18.90; 1.12	74%	Students	FFNI-SF (NI)	FFNI-SF (NI)	Family, friends	–	0.38	–	–	75.00
Bernerth et al. (2022)	United States	291	NI	NI	Leaders	DD (0.83)	DD (0.72)	Subordinates	–	0.20	–	–	66.07
	United States	98	NI	34%	Leaders	DD (0.77)	DD (0.74)	Subordinates	–	0.38	–	–	
Kardum et al. (2022)	Croatia	188	39.55; 11.71	100%	General population	SRP-III (0.76); NPI-40 (0.80); Mach-IV (0.71)	SRP-III (0.83); NPI-40 (0.83); Mach-IV (0.70)	Romantic partners	0.53	0.60	0.63	–	64.29
	Croatia	188	42.15; 12.13	0%	General population	SRP-III (0.81); NPI-40 (0.89); Mach-IV (0.71)	SRP-III (0.87); NPI-40 (0.90); Mach-IV (0.80)	Romantic partners	0.64	0.52	0.53	–	

ASSESSING THE DARK PERSONALITY TRAITS WITH OBSERVER REPORTS: A META-ANALYSIS OF INTER-RATER AGREEMENT ON THE DARK TRIAD AND DARK TETRAD TRAITS

7

TABLE 1 (Continued)

Study	Observation samples					Instruments			Results (r)				STROBE score
	Country	Size (N)	Sample age (M; SD)	Gender proportion (female)	Sample type	Self-report (α or ω)	Observer report (α or ω)	Informant	P	N	M	S	
Niemeyer et al. (2022)	Germany	199	37.56; 12.65	0%	Prison inmates	SD3 (0.71); SD3 (0.61); SD3 (0.71)	SD3 (0.85); SD3 (0.89); SD3 (0.87)	Prison workers	0.44	0.36	0.32	NI	75.00
Uzieblo et al. (2022)	Belgium	258	32.20; NI	0%	General population and students	SRP-SF (0.86)	SRP-SF (0.86)	Romantic partners	0.49	–	–	–	67.86
Kardum, Hudek-Knezevic, Mehić, and Banov Trošelj (2023)	Croatia	205	27.17; 5.06	100%	General population	SRP-III (0.81); NPI-40 (0.84); Mach-IV (0.72)	SRP-III (0.84); NPI-40 (0.87); Mach-IV (0.71)	Romantic partners	0.56	0.57	0.44	–	80.36
	Croatia	205	29.40; 6.48	0%	General population	SRP-III (0.81); NPI-40 (0.87); Mach-IV (0.75)	SRP-III (0.88); NPI-40 (0.91); Mach-IV (0.80)	Romantic partners	0.64	0.59	0.51	–	
Kardum, Hudek-Knezevic, Marijanović, and Shackelford (2023)	Croatia	187	23.40; 3.29	100%	General population	DD (0.67); DD (0.88); DD (0.82)	DD (0.66); DD (0.91); DD (0.86)	Romantic partners	0.32	0.37	0.37	NI	66.07
	Croatia	187	24.91; 3.62	0%	General population	DD (0.63); DD (0.85); DD (0.83)	DD (0.61); DD (0.85); DD (0.84)	Romantic partners	0.42	0.48	0.54	NI	
Pfattheicher et al. (2023)	Serbia	314	18.03; 0.34	55.10%	Students	CAST (0.89)	CAST (0.93)	Classmate	–	–	–	0.36	85.71
Szabó et al. (2023)	Germany and Hungary	91	43.62; 10.31	46.15%	Leaders	SD3 (0.71); SD3 (0.66); SD3 (0.76)	SD3 (0.74); SD3 (0.63); SD3 (0.79)	Subordinates	0.26	0.44	0.34	NI	71.43
Walker et al. (2023)	NI	411	36.70; 10.19	43.80%	Prolific	DD (0.70); DD (0.80); DD (0.81)	DIRT (0.78); DIRT (0.85); DIRT (0.85)	Romantic partners	0.47	0.38	0.41	NI	75.00

Abbreviations: Brief DT, Brief Triad Scale; DD, Dirty Dozen; DD-Y, Dirty Dozen for Youths; DIRT, Dark Informant-Rated Triad; FFNI-SF, Five-Factor Narcissism Inventory-Short Form; IM-N, Interpersonal Measure of Narcissism; LSRP, Levenson Self-Report Psychopathy Scale; M, Machiavellianism; Mach-IV, Machiavellianism Scale; Mturk, Amazon's Mechanical Turk; N, narcissism; NEO-FFI, NEO Five-Factor Inventory; NI, not indicated; NPI, Narcissistic Personality Inventory; P, psychopathy; r, Pearson's correlation; S, sadism; SD3, Short Dark Triad; SRP-III, Self-Report Psychopathy Scale-III version; SRP-SF, Self-Report Psychopathy-Short Form; CAST, Comprehensive Assessment of Sadistic Tendencies; STROBE, Strengthening the Reporting of Observational Studies in Epidemiology; α, Cronbach's alpha; ω, McDonald's omega.

[a]By Fisher's z transformation.

(approximately 10% of all studies), and subsequently, one reviewer analysed the rest of the studies. To obtain an index of inter-rater reliability, a correlational analysis was performed, which indicated good agreement: $r = 0.85$, $p < 0.001$.

Statistical analysis

To estimate the magnitude of the relationships between the study variables (i.e., the Dark Tetrad traits assessed by self-report and observer report), effect sizes were calculated from the correlations (Pearson's r) provided in the studies included in the meta-analysis. To interpret the magnitudes of the correlations, the criteria established by Cohen (1992) were used as a reference: A value of $r = 0.10$ or close to it indicates a small magnitude, a value of $r = 0.30$ or close to it indicates a medium magnitude and a value of $r = 0.50$ or higher indicates a large magnitude.

To include the type of informant as a moderator, this variable had to be categorized. For this purpose, two analyses were tested. In the first, the variable was dichotomized into $0 = $ *very/somewhat acquainted* and $1 = $ *not at all acquainted*, and in the second, it was trichotomized into $0 = $ *very well acquainted*, $1 = $ *somewhat acquainted* and $2 = $ *not at all acquainted*. This decision was made because, in the studies, we found informants ranging from very well *acquainted*, such as parents, to not at all *acquainted*, such as Facebook users, and informants who were somewhat *acquainted*, such as subordinates. In this sense, following the literature, we wanted to test possible differences between the two ways of categorizing this variable. To interpret the results, both the intercept and the moderator effect were included. The intercept represents the estimated effect size when the moderator is set to the reference category (e.g., *very/somewhat acquainted* in the dichotomous model). The moderator effect coefficient indicates the change in the effect size when the moderator shifts from the reference category to the other categories.

One of the studies (Lämmle et al., 2021) provided three correlations between self-report and observer report for each trait, so, to obtain a single correlational value and not bias the results (and avoid duplicating the sample), the r values were transformed into Fisher's z, averaged to obtain a single value, and transformed back into r values (Sánchez-Bruno & Borges del Rosal, 2005).

The analysis was carried out using Fisher's r-to-z transformed correlation coefficient as the outcome measure, and the data were analysed using a random-effects model. The restricted maximum likelihood estimator (REML; Viechtbauer, 2010) was utilized to estimate the amount of heterogeneity (i.e., τ^2) in the data. Furthermore, we assessed heterogeneity using the Q-test (Cochran, 1954) and calculated Higgins' test (I^2), as well as 95% prediction intervals for the observed results.

Finally, publication biases were examined using the Begg and Mazumdar (1994) rank correlation test, two tests that assess possible effects due to the inclusion of small studies. This test calculates a correlation coefficient between the effect size and its variance in a funnel plot, and a significant correlation (p-value < 0.05) suggests the presence of effects. However, this test may be underpowered to detect effect when the number of studies is small. Therefore, Egger et al.'s (1997) regression test was also employed to increase sensitivity in detecting the effect. This test examines the relationship between the effect size and its precision (standard error) in a funnel plot and performs a linear regression analysis. Likewise, a significant intercept (p-value <0.05) suggests the presence of effects. All statistical analyses were performed using the Jamovi (version 2.2.5) statistical program.

RESULTS

Selection and inclusion of studies

Figure 1 shows the process of searching, screening and excluding studies to obtain the final number of studies included in the meta-analysis ($N = 24$ with 32 effect sizes). Thus, a total of 12,678 studies were obtained after adding the search equation to the different databases. Of these, 5991 were eliminated as duplicates, resulting in a total of 6688 studies to be screened. In the first screening, most of them were excluded because they were not related to the topic of interest (i.e., they included, for example, the terms 'triad', 'tetrad' or 'dark' but did not refer to the Dark Triad or Dark Tetrad). It was not possible to discard any studies due to exclusion reason 8 (i.e., 'does not use an observer report to assess the Dark Triad or Dark Tetrad traits') because, for most studies, we could not determine the inclusion of peer reviews, so this reason turned out to be unique to the second screening. In the second screening, 2404 studies were excluded mainly for not employing observer reports and self-reports simultaneously for assessing dark traits.

Quality assessment of studies

The methodological quality of the selected studies was evaluated, revealing that only one study was rated as excellent, but none was rated as poor, and that 12 studies were rated as good, and 11 studies were deemed fair.

Most of the studies were incomplete in providing comprehensive information on their research methodology, specifically concerning details on the sampling strategy and statistical approaches. However, there was no moderation effect considering the quality of the studies in the values of interest.

Characteristics of the included studies

General information

Table 1 presents the characteristics of the 24 papers that met the established inclusion criteria. However, although 24 papers were included, eight of them reported two samples because they split the analysis between independent informants or between men and women, because they used more than one self-report instrument or because they conducted two independent studies. For this reason, they were considered two independent studies when performing the analyses. As a result, 24 studies with 32 effect sizes were included in the meta-analysis.

The papers were published between 2013 and 2023, and only one of the studies assessed the Dark Tetra (i.e., only one included the assessment of sadism). Therefore, 23 studies assessed narcissism, Machiavellianism and psychopathy (the Dark Triad), so the results presented in this section correspond to this set of three traits (it was not possible to include sadism in the analyses).

Description of the sample

The total study sample (i.e., the total number of participants evaluated by both self-report and observer report in all the included studies) was 7022 (mean sample size = 219.44; range = 62 to 411). Concerning their characteristics, the mean age across samples was 26.41, and the mean proportion of women was 52%. Most of the participants were Americans (approximately 36.12%) or Germans (approximately 22.09%), and students (approximately 47.78%) or general population (approximately 38.56%).

Description of self-report and observer report scales

Nine different types of scales were used (in their original or most recent versions), including the Self-Report Psychopathy Scale (SRP; Paulhus et al., 2009), Levenson Self-Report Psychopathy Scale (LSRP; Levenson et al., 1995) the Narcissistic Personality Inventory (NPI;

Raskin & Hall, 1979), the Five-Factor Narcissism Inventory-Short Form (FFNI-SF; Sherman et al., 2015), the Machiavellianism Scale (Mach-IV; Christie & Geis, 1970), the Short Dark Triad (SD3; Jones & Paulhus, 2014), the Dirty Dozen (DD; Jonason & Webster, 2010), a combination of the SD3 and a brief Dark Triad measure, and an ad hoc version of the NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992). The internal consistency estimates (Cronbach's alpha—α or McDonald's omega—ω) for these instruments ranged from 0.43 to 0.93. The same scales were used in each study to collect observer reports, modifying the items to the third person. However, four of them used different scales: an ad hoc scale, a combination of the SD3 and DD, the Interpersonal Measure of Narcissism (IM-N; an unpublished measure) and a validation of the Dark Informant-Rated Triad scale (DIRT; derived from the DD). In this case, the internal consistency estimates ranged from −0.04 to 0.98. The informants included family members, friends, acquaintances from high school, college, university or hometown, roommates, former or current romantic partners, trained research assistants, prison workers and subordinates. Some studies selected informants unknown to the targets (i.e., Facebook people and trained research assistants).

Self-observer correlations for narcissism

The random model ($k = 26$) showed a positive self-observer correlation of medium magnitude at the intercept ($r = 0.44$, 95% CI [0.35, 0.52], $z = 10.50$, $p < 0.001$). The effect size estimates were obtained using Fisher's r-to-z transformation to stabilize the variance of the correlations, and, after conducting the analysis, the results were transformed back to Pearson's r for ease of interpretation. The Q-test and I^2 statistics indicated significant heterogeneity between samples, Q (25) = 215.43, $p < 0.001$, $I^2 = 89.16$. The estimate of τ^2 was 0.04, suggesting variability in effect sizes across studies. The 95% prediction interval, which provides an estimate of the range where future study effects are likely to fall, was 0.04–0.83. The forest plot of the results is shown in Figure 2.

When informant type was included in the analysis, the random model ($k = 26$) yielded a significant positive intercept of $r = 0.64$ (95% CI [0.34, 0.94], $p < 0.001$) when informant type was dichotomized (0 = *very well-somewhat acquainted*; 1 = *not at all acquainted*), and a negative correlation of medium magnitude for the moderator effect ($r = -0.19$, 95% CI [−0.44, 0.07], $p > 0.05$). When informant type was trichotomized (0 = *very well acquainted*; 1 = *somewhat acquainted*; 2 = *not at*
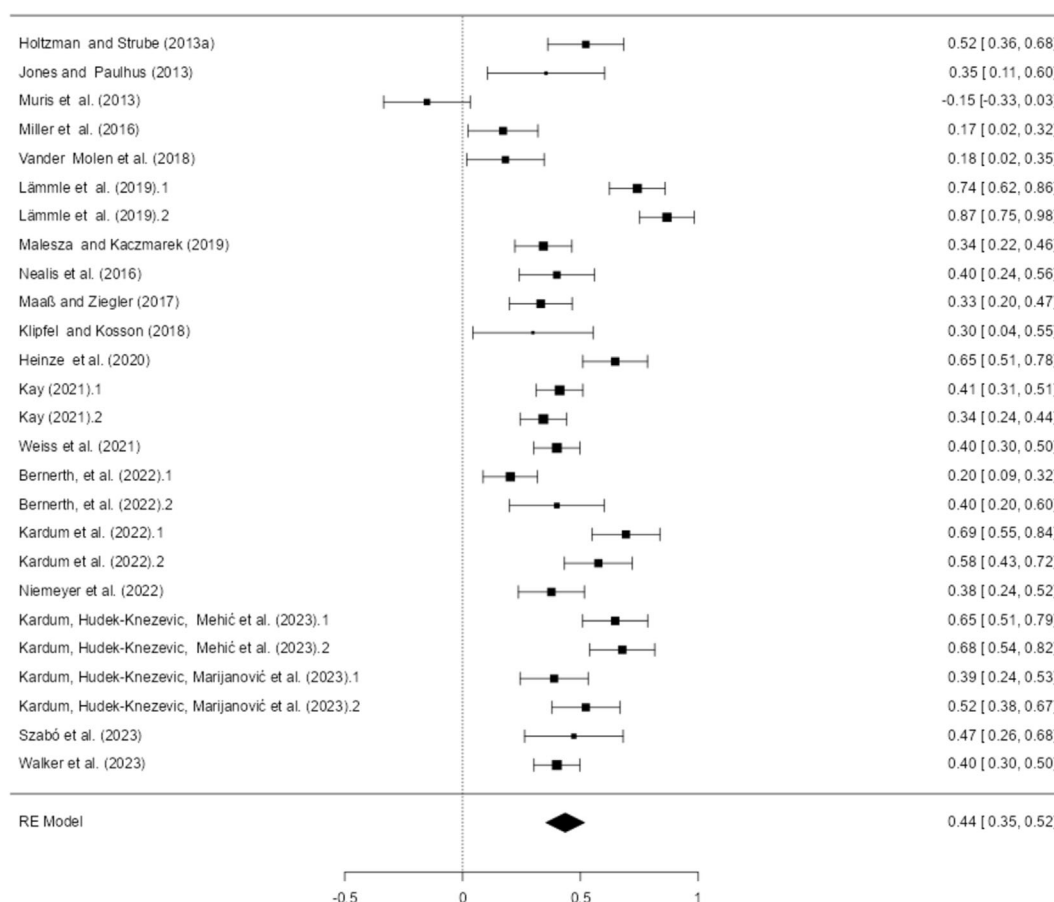
Holtzman and Strube (2013a)	0.52 [0.36, 0.68]
Jones and Paulhus (2013)	0.35 [0.11, 0.60]
Muris et al. (2013)	−0.15 [−0.33, 0.03]
Miller et al. (2016)	0.17 [0.02, 0.32]
Vander Molen et al. (2018)	0.18 [0.02, 0.35]
Lämmle et al. (2019).1	0.74 [0.62, 0.86]
Lämmle et al. (2019).2	0.87 [0.75, 0.98]
Malesza and Kaczmarek (2019)	0.34 [0.22, 0.46]
Nealis et al. (2016)	0.40 [0.24, 0.56]
Maaß and Ziegler (2017)	0.33 [0.20, 0.47]
Klipfel and Kosson (2018)	0.30 [0.04, 0.55]
Heinze et al. (2020)	0.65 [0.51, 0.78]
Kay (2021).1	0.41 [0.31, 0.51]
Kay (2021).2	0.34 [0.24, 0.44]
Weiss et al. (2021)	0.40 [0.30, 0.50]
Bernerth, et al. (2022).1	0.20 [0.09, 0.32]
Bernerth, et al. (2022).2	0.40 [0.20, 0.60]
Kardum et al. (2022).1	0.69 [0.55, 0.84]
Kardum et al. (2022).2	0.58 [0.43, 0.72]
Niemeyer et al. (2022)	0.38 [0.24, 0.52]
Kardum, Hudek-Knezevic, Mehić et al. (2023).1	0.65 [0.51, 0.79]
Kardum, Hudek-Knezevic, Mehić et al. (2023).2	0.68 [0.54, 0.82]
Kardum, Hudek-Knezevic, Marijanović et al. (2023).1	0.39 [0.24, 0.53]
Kardum, Hudek-Knezevic, Marijanović et al. (2023).2	0.52 [0.38, 0.67]
Szabó et al. (2023)	0.47 [0.26, 0.68]
Walker et al. (2023)	0.40 [0.30, 0.50]
RE Model	0.44 [0.35, 0.52]

FIGURE 2   Forest plot of the relationship between narcissism assessed with self-reports and observer reports. RE, random effects.

acquainted), the intercept indicated a positive correlation of $r = 0.58$ (95% CI [0.40, 0.76], $p < 0.001$), and the moderator effect yielded a negative correlation of small magnitude ($r = -0.11$, 95% CI [−0.22, 0.01], $p > 0.05$). These results indicated that the informant type is not a significant moderator of the relationship between the two assessment methodologies.

Self-observer correlations for Machiavellianism

The random-effects model ($k = 19$) yielded a positive medium magnitude correlation at the intercept ($r = 0.41$, 95% CI [0.34, 0.47], $z = 11.70$, $p < 0.001$) between self-reports and observer reports. Effect size estimates were also obtained by Fisher's r-to-z transformation. The Q-test and I^2 statistics indicated significant heterogeneity between samples, $Q (18) = 77.38$, $p < 0.001$, $I^2 = 77.64$. The estimate of τ^2 was 0.02, suggesting variability in effect sizes across studies. The 95% prediction interval was 0.14–0.67. The forest plot is presented in Figure 3.

In this case, the random model ($k = 19$) yielded a significant positive intercept of $r = 0.72$ (95% CI [0.41, 1.04], $p < 0.001$) when informant type was dichotomized

($0 = $ *very well-somewhat acquainted*; $1 = $ *not at all acquainted*), and a negative correlation of medium magnitude for the moderator effect ($r = -0.30$, 95% CI [−0.59, −0.01], $p < 0.05$). When informant type was trichotomized ($0 = $ *very well acquainted*; $1 = $ *somewhat acquainted*; $2 = $ *not at all acquainted*), the intercept indicated a positive correlation of $r = 0.57$ (95% CI [0.41, 0.73], $p < 0.001$), and the moderator effect yielded a negative correlation of small magnitude ($r = -0.14$, 95% CI [−0.26, −0.01], $p < 0.05$). These results indicated that the informant type is a significant moderator of the relationship between the two assessment methodologies and that this relationship will be stronger when the informant is someone known to the informant.

Self-observer correlations for psychopathy

Finally, the random-effects model ($k = 22$) produced a positive self-observer correlation of medium-high magnitude at the intercept ($r = 0.47$, 95% CI [0.40, 0.53], $z = 13.70$, $p < 0.001$). Effect size estimates were also obtained by Fisher's r-to-z transformation. The Q-test and I^2 statistics indicated significant heterogeneity between
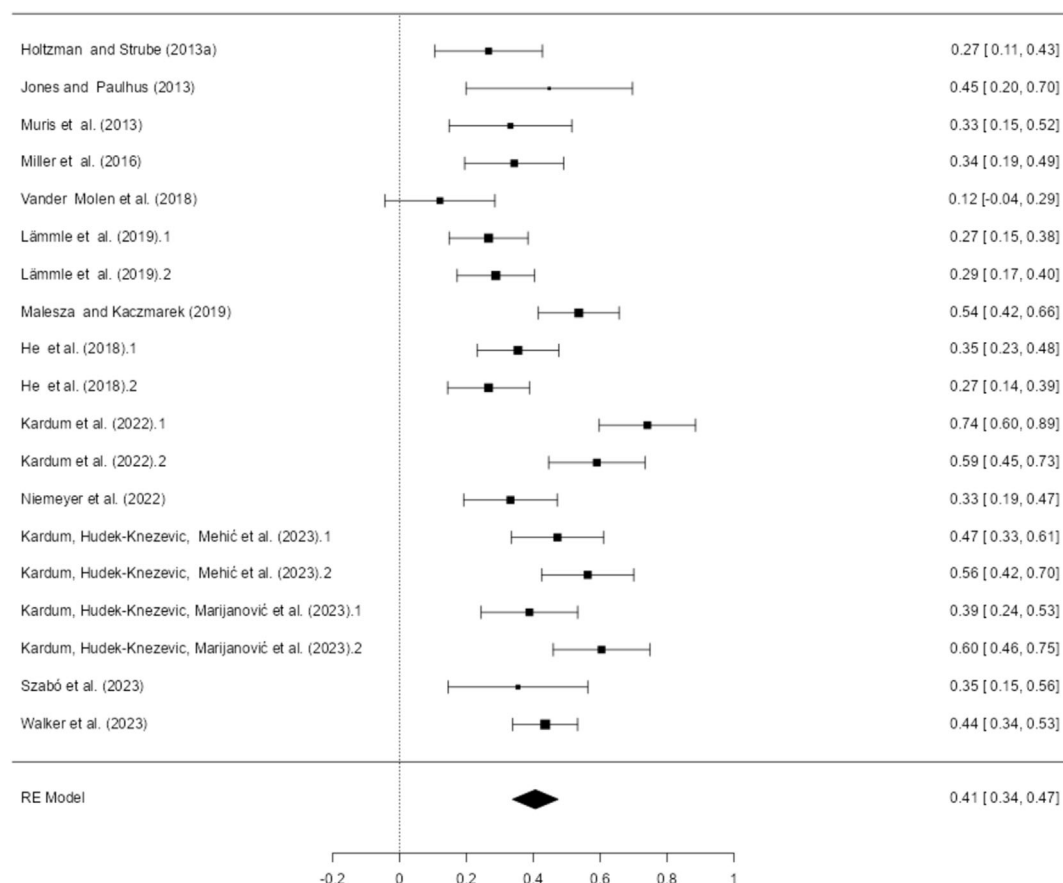
FIGURE 3 Forest plot of the relationship between Machiavellianism assessed with self-reports and observer reports. RE, random effects.

samples, Q (21) $= 92.39$, $p < 0.001$, $I^2 = 79.78$. The estimate of τ^2 was 0.02, suggesting variability in effect sizes across studies. The 95% prediction interval was 0.18–0.75. The forest plot is presented in Figure 4.

The random model ($k = 22$) yielded a significant positive intercept of $r = 0.88$ (95% CI [0.57, 1.19], $p < 0.001$) when informant type was dichotomized (0 = *very well-somewhat acquainted*; 1 = *not at all acquainted*), and a negative correlation of medium magnitude for the moderator effect ($r = -0.39$, 95% CI [-0.69, -0.10], $p < 0.01$). When informant type was trichotomized (0 = *very well acquainted*; 1 = *somewhat acquainted*; 2 = *not at all acquainted*), the intercept indicated a positive correlation of $r = 0.67$ (95% CI [0.52, 0.83], $p < 0.001$), and the moderator effect yielded a negative correlation of small magnitude ($r = -0.17$, 95% CI [-0.30, -0.05], $p < 0.01$). These results were similar to those for Machiavellianism.

Publication bias—small study effects

Possible effect of small studies was examined using Begg and Mazumdar's rank correlation and Egger's regression test. The Begg and Mazumdar rank correlation test

yielded the following results: for narcissism, a correlation coefficient of -0.08 ($p = 0.58$); for Machiavellianism, a correlation of 0.07 ($p = 0.70$); and for psychopathy, a correlation of -0.32 ($p = 0.04$). The Egger's regression test showed the following results: for narcissism, a regression coefficient of -1.22 ($p = 0.22$); for Machiavellianism, a coefficient of -0.19 ($p = 0.85$); and for psychopathy, a coefficient of -1.08 ($p = 0.28$). Although there is an indication of publication bias in the case of psychopathy according to Begg and Mazumdar's test, there is no strong confirmation due to the lack of significance in Egger's test. Therefore, it can be extracted that there is no evidence of effects due to the inclusion of small studies in any of the cases analysed.

DISCUSSION

Results obtained in the analyses of this study revealed positive correlations of medium magnitude between the three traits assessed with self-reports and observer reports, indicating a positive relationship between these measures. In the case of psychopathy, although it does not exceed 0.50, its value is closer to 0.50 than to 0.30, so
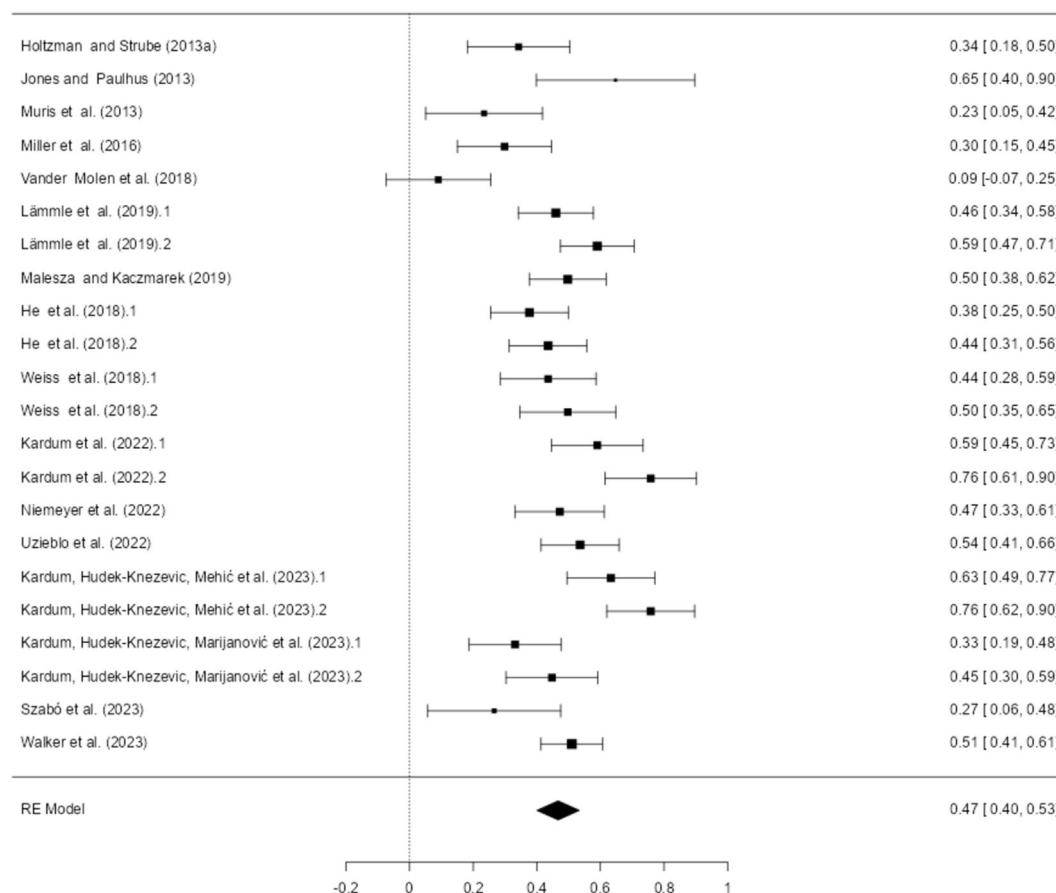
Study		Estimate
Holtzman and Strube (2013a)		0.34 [0.18, 0.50]
Jones and Paulhus (2013)		0.65 [0.40, 0.90]
Muris et al. (2013)		0.23 [0.05, 0.42]
Miller et al. (2016)		0.30 [0.15, 0.45]
Vander Molen et al. (2018)		0.09 [-0.07, 0.25]
Lämmle et al. (2019).1		0.46 [0.34, 0.58]
Lämmle et al. (2019).2		0.59 [0.47, 0.71]
Malesza and Kaczmarek (2019)		0.50 [0.38, 0.62]
He et al. (2018).1		0.38 [0.25, 0.50]
He et al. (2018).2		0.44 [0.31, 0.56]
Weiss et al. (2018).1		0.44 [0.28, 0.59]
Weiss et al. (2018).2		0.50 [0.35, 0.65]
Kardum et al. (2022).1		0.59 [0.45, 0.73]
Kardum et al. (2022).2		0.76 [0.61, 0.90]
Niemeyer et al. (2022)		0.47 [0.33, 0.61]
Uzieblo et al. (2022)		0.54 [0.41, 0.66]
Kardum, Hudek-Knezevic, Mehić et al. (2023).1		0.63 [0.49, 0.77]
Kardum, Hudek-Knezevic, Mehić et al. (2023).2		0.76 [0.62, 0.90]
Kardum, Hudek-Knezevic, Marijanović et al. (2023).1		0.33 [0.19, 0.48]
Kardum, Hudek-Knezevic, Marijanović et al. (2023).2		0.45 [0.30, 0.59]
Szabó et al. (2023)		0.27 [0.06, 0.48]
Walker et al. (2023)		0.51 [0.41, 0.61]
RE Model		0.47 [0.40, 0.53]

FIGURE 4   Forest plot of the relationship between psychopathy assessed with self-reports and observer reports. RE, random effects.

we considered it a large magnitude association. Thus, the correlations obtained between self and observer could indicate that the assessments given by observers will closely match the targets' perceptions of their own personality (Jones & Paulhus, 2014; Luan et al., 2019; Malesza & Kaczmarek, 2020; Vazire, 2010). In line with the studies of Malesza and Kaczmarek (2020) and Mischel (1968), these results provide support for the convergence between self-reports and observer reports because, for all three traits, the correlations exceed the 0.30 validity cutoff.

Not only can we explain these moderate correlations by the raters' accuracy (alternative assessment procedures for the same construct) but also by the fact that the raters might be observing distinct but related phenomena using different samples of behaviours to reach a conclusion about the same trait (Larsen et al., 2017). For example, according to previous research findings, a possible reason for the ease with which narcissism is observed is that familiarity tends to attract others in similar circumstances to the targets. This means that narcissism exhibits a consistent pattern with different people. In contrast, it is argued that, when examining Machiavellianism and psychopathy, raters may be using less similar information

or interpreting it differently (Lämmle et al., 2021; Maaß & Ziegler, 2017). However, the results of our meta-analysis show that psychopathy appears to be one of the most easily observable traits as observers tended to agree with the scores obtained in the self-reports (e.g., He et al., 2018; Jones & Paulhus, 2014; Malesza & Kaczmarek, 2020).

In this sense, people with narcissistic traits seem to be aware of their personality and the need to maintain a narcissistic reputation, so they tend to show themselves to others as they are (i.e., narcissistic). In other words, their need for public recognition may lead them to exhibit behaviour that tends to attract the attention and admiration of others, making it an easily observable trait. In contrast, people with Machiavellian traits (the least observable trait, according to the results obtained) might present themselves to others as 'good' in certain situations, being able to deceive and manipulate others without this being perceived. This hidden nature of Machiavellianism makes more difficult for others to observe this trait in people, leading to an underestimation of the trait (Carlson et al., 2011; Lämmle et al., 2021; Maaß & Ziegler, 2017). In the case of people with psychopathic traits, they may also present themselves as they are, like narcissists, and show a lack of concern for the

consequences of how they appear to others (e.g., Miller et al., 2011). In fact, this differentiation between the visibility of traits is in line with the definition of each trait. Hence, the subtle variations in effect sizes observed in this meta-analysis may stem from the inherent nature of the traits (Paulhus & Williams, 2002). However, although the included studies reach different conclusions as to which of the three traits is easier to observe, this meta-analysis proved that all three traits are observed with approximately the same ease.

According to the meta-analysis conducted by Connelly and Ones (2010), there is a moderate correlation between self-reported personality traits and other-reported personality traits. Specifically, the meta-analysis found correlation coefficients between 0.08 and 0.48, with the highest correlations appearing when the others were people closer to the target. This indicated a positive relationship between the two types of personality reports, which was later corroborated in the meta-analysis of Kim et al. (2019), who found that the means of self-reports generally did not differ from the means of informant reports (mean $\delta = -0.04$). Furthermore, the former meta-analysis found that the accuracy of observer-reported personality traits varied depending on the trait being assessed. For example, observers were found to be more accurate in assessing extraversion and openness to experience, but less accurate in assessing neuroticism and agreeableness, arguing that observer reports might be more useful for highly observable traits (Luan et al., 2019; Vazire, 2010).

Although no conclusive statements can be drawn due to the difference in sample sizes between our meta-analysis and that of Connelly and Ones (2010), it is noteworthy that the effect sizes obtained in our analysis of dark traits are higher than those obtained in the analysis of the Big Five traits (around 0.40 versus 0.30, except in the case of extraversion). As already argued, this could be due to the fact that socially highly desirable or highly undesirable traits, such as dark traits, are more easily observed by others (Luan et al., 2019). Moreover, Funder and Dobroth (1987) previously examined the varying observability of personality traits. They established that a trait is more readily observable when its associated and unassociated behaviours are easier to conceptualize, there are ample opportunities for such behaviours, only a few affirmative behaviours are needed to define the trait and its assessment appears intuitively straightforward. However, more recent studies indicate that it is difficult to establish differences in observability, with some traits being easier to observe than others (e.g., Allik et al., 2016).

In any case, we also note that in the studies included in our meta-analysis, the highest correlations are those where the raters of the dark personality were acquaintances of the targets (such as family or friends) (Maaß & Ziegler, 2017; Vander Molen et al., 2018). Thus, our study showed that informant type is a moderating variable of the relationship between traits measured with self-report and with observer report. However, these results were only found for Machiavellianism and psychopathy; no such significant moderation was obtained for narcissism. Specifically, our results suggest that Machiavellianism and psychopathy are more easily observable by those with a closer relationship with the target, so the greater the familiarity, the greater the accuracy in trait assessment (Connelly & Ones, 2010; Vazire, 2006; Vazire & Mehl, 2008). These findings are in line with those obtained in previous studies in which higher correlations were found for Big Five traits when the observer knew the person being assessed (Connelly & Ones, 2010; Connolly et al., 2007; Kim et al., 2019). In the case of narcissism, a possible explanation for the results found is the one provided by Szabó et al. (2023) in their study of leaders and subordinates. In their study, it was shown that, unlike Machiavellianism and psychopathy, narcissism is generally positively associated with achievement motivation, status orientation, need for identity and organizational identification, among others (Prusik & Szulawski; 2019). These characteristics demonstrate, again, that narcissistic people behave in the same way, that is, narcissistic, regardless of who they are with (Lämmle et al., 2021; Maaß & Ziegler, 2017). Thus, a narcissistic person may be easier to identify, regardless of whether they are being observed by a total stranger (such as a Facebook user; Vander Molen et al., 2018) or a person who does not know them as well as a romantic partner would (as could be a subordinate; Szabó et al., 2023).

On the other hand, it is important to mention that only one study was located that used observer reports to measure sadism. A reason for this may be that this trait was included later in the set of dark traits (Chabrol et al., 2009). In addition, the more recent development of self-report measures specifically designed to assess this trait along with the other three traits could also have contributed to this lack of studies (Paulhus et al., 2021; Plouffe et al., 2017; Webster & Wongsomboon, 2020).

It is also interesting to note the infrequency with which this methodology has been used to assess these dark traits. This fact may be due to the existing debate in the literature on their accuracy and the disadvantages that this methodology may present. Some of these disadvantages are the additional effort required to obtain data from third parties and that there is some personal information that informants cannot access. Other

disadvantages are the possible difficulty in assessing very specific behaviours, and the possible presence of biases similar to those of self-reports, such as extreme responses or acquiescence (Baker et al., 2004; Malesza & Kaczmarek, 2020; McDonald, 2008).

In any case, when interpreting the results obtained in this study, it is underlined that a significant heterogeneity was observed between samples. This indicates that the correlation between self-reports and observer reports may vary depending on the sample characteristics, including differences in the sample sizes of the included studies. Furthermore, the significant τ^2 values obtained for each analysis suggest that there may be unexplained sources of heterogeneity that should be explored in future research (Hoaglin, 2016).

Finally, it is also important to note that our findings are consistent with previous research that has explored the agreement between self and observer reports in the assessment of personality disorder traits, such as antisocial (Oltmanns & Oltmanns, 2021; Oltmanns & Widiger, 2021; Tyrer et al., 2021). These studies (one of which is a meta-analysis) have also demonstrated moderate correlations, which reinforces the validity of using observers to assess antagonistic traits and which in turn adds robustness to our conclusions about Dark Tetrad traits.

Limitations and future research

The present study presents some limitations. First, not many studies met the inclusion criteria, so the analyses were performed with a small sample of studies, which makes it difficult to generalize the results. In addition, the design of these studies was cross-sectional and not longitudinal, which, again, hinders generalizing the results. As a future line of research, it is proposed to continue updating this meta-analysis to provide more reliable results on the correlation between self and observer reports. Likewise, with a larger sample size and an expanded meta-analysis, consideration could also be given to examining other possible moderators, guided by previous literature and thus updating knowledge on observer reports.

Finally, we note that the validity of some of the short scales used in the studies included in this meta-analysis (i.e., the SD3 and the DD) has been questioned, considering that these shorter measures may not assess all aspects of the dark triad traits. Therefore, it is of great importance to review the psychometric properties of both self-reported and hetero-reported scales when conducting research (Muris et al., 2017).

Conclusions

This is the first meta-analysis of the studies that have employed observer reports to measure Dark Traits and provides statistical data on the correlation between self and observer reports. Evidence is provided that observer report could be a good method for indirectly assessing dark traits and, therefore, offset and complement a self-report assessment (Abernethy, 2015; Connelly & Ones, 2010; Kholin et al., 2016; Kim et al., 2019; Luan et al., 2019; Malesza & Kaczmarek, 2020; Müller & Moshagen, 2019; Muris et al., 2017; Vazire, 2010; Vazire & Carlson, 2011). The basis of this work was that observer reports have been found to be especially useful when dealing with socially desirable or undesirable (rather than neutral) traits because their self-assessment presents greater biases of presenting oneself in more socially acceptable terms (Andrews & Meyer, 2003; Echeburúa et al., 2011; Luan et al., 2019; Spaans et al., 2017). In addition, a comprehensive and robust assessment of any given construct is more effectively achieved by integrating multiple measurement methods (e.g., Kyllonen & Kell, 2018; McDonald, 2008; Ortner & Proyer, 2015).

Given the association between these personality traits and numerous forms of uncivil behaviour (such as lack of dishonesty-humility, physical, verbal or sexual aggressions, cyberviolence, bullying and sextortion, among others), it is crucial to take into account potential response biases related to social desirability and to utilize a diverse range of assessment methods to enhance the accuracy and reliability of findings (Alsheikh Ali, 2020; Hayes et al., 2021; Moor & Anderson, 2019; Pineda et al., 2023; Pineda, Galán, et al., 2022; Pineda, Rico-Bordera, et al., 2022; Rico-Bordera et al., 2023; Rogier et al., 2021; Roters & Book, 2023; Spaans et al., 2017; Thomas & Egan, 2022). Incorporating observer reports can aid in the identification and understanding of dark personality traits, which may be particularly relevant for interventions targeting maladaptive behaviours associated with these traits.

CONFLICT OF INTEREST STATEMENT
The authors declare that there is no conflict of interest.

DATA AVAILABILITY STATEMENT
All data generated or analysed during this study are included in this published article (Table 1). The analysis scripts used for this article and information on the origins of the data presented in Table 1 can be accessed at https://osf.io/fbjrt/?view_only=75091fbbfd01403283096e3fea2f4b92.

ORCID
Pilar Rico-Bordera https://orcid.org/0000-0002-2157-2468
David Pineda https://orcid.org/0000-0001-5710-1636
Manuel Galán https://orcid.org/0000-0003-4730-0695
José A. Piqueras https://orcid.org/0000-0002-3604-5441

REFERENCES
Abernethy, M. (2015). Self-reports and observer reports as data generation methods: An assessment of issues of both methods. *Universal Journal of Psychology*, *3*(1), 22–27. https://doi.org/10.13189/ujp.2015.030104

Allik, J., de Vries, R. E., & Realo, A. (2016). Why are moderators of self-other agreement difficult to establish? *Journal of Research in Personality*, *63*, 72–83. https://doi.org/10.1016/j.jrp.2016.05.013

Alsheikh Ali, A. S. (2020). Delinquency as predicted by dark triad factors and demographic variables. *International Journal of Adolescence and Youth*, *25*(1), 661–675. https://doi.org/10.1080/02673843.2020.1711784

Andrews, P., & Meyer, R. G. (2003). Marlowe–Crowne Social Desirability Scale and short form C: Forensic norms. *Journal of Clinical Psychology*, *59*(4), 483–492. https://doi.org/10.1002/jclp.10136

Baker, S. R., Victor, J. B., Chambers, A. L., & Halverson, C. F. (2004). Adolescent personality: A five-factor model construct validation. *Assessment*, *11*(4), 303–315. https://doi.org/10.1177/1073191104269871

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, *50*(4), 1088–1101. https://doi.org/10.2307/2533446

Bernerth, J. B., Carter, M. Z., & Cole, M. S. (2022). The (in)congruence effect of leaders' narcissism identity and reputation on performance: A socioanalytic multistakeholder perspective. *Journal of Applied Psychology*, *107*(10), 1725–1742. https://doi.org/10.1037/apl0000974

Carlson, E. N., Vazire, S., & Oltmanns, T. F. (2011). You probably think this paper's about you: Narcissists' perceptions of their personality and reputation. *Journal of Personality and Social Psychology*, *101*, 185–201. https://doi.org/10.1037/a0023781

Cattell, R. B., & Warburton, F. W. (1967). *Objective personality and motivation tests: A theoretical introduction and practical compendium*. University of Illinois Press.

Chabrol, H., Van Leeuwen, N., Rodgers, R., & Séjourné, N. (2009). Contributions of psychopathic, narcissistic, Machiavellian, and sadistic personality traits to juvenile delinquency. *Personality and Individual Differences*, *47*(7), 734–739. https://doi.org/10.1016/j.paid.2009.06.020

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

Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics*, *10*(1), 101–129. https://doi.org/10.2307/3001666

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Connelly, B. S., & Ones, D. S. (2010). Another perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, *136*(6), 1092–1122. https://doi.org/10.1037/a0021212

Connolly, J. J., Kavanagh, E. J., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International Journal of Selection and Assessment*, *15*(1), 110–117. https://doi.org/10.1111/j.1468-2389.2007.00371.x

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEP Five-Factor Inventory (NEO-FFI): Professional manual*. Psychological Assessment Resources.

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*, *3*(2), 117–131. https://doi.org/10.1365/s42681-021-00025-6

Echeburúa, E., Muñoz, J. M., & Loinaz, I. (2011). La evaluación psicológica forense frente a la evaluación clínica: Propuestas y retos de futuro. [Forensic psychological assessment versus clinical assessment: Proposals and challenges for the future]. *International Journal of Clinical and Health Psychology*, *11*(1), 141–159. https://www.redalyc.org/pdf/337/33715423009.pdf

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal (Clinical Research Edition)*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, *102*(4), 652–670. https://doi.org/10.1037/0033-295X.102.4.652

Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology*, *52*(2), 409–418. https://doi.org/10.1037//0022-3514.52.2.409

Hare, R. D. (1970). *Psychopathy: Theory research, and implications for society*. Wiley.

Hayes, N. L., Marsee, M. A., & Russell, D. W. (2021). Latent profile analysis of traditional and cyber-aggression and victimization: Associations with dark triad traits and psychopathology symptoms. *Journal of Psychopathology and Behavioral Assessment*, *43*(2), 399–412. https://doi.org/10.1007/s10862-020-09835-2

He, Q., Wang, Y., Xing, Y., & Yu, Y. (2018). Dark personality, interpersonal rejection, and marital stability of Chinese couples: An actor–partner interdependence mediation model. *Personality and Individual Differences*, *134*, 232–238. https://doi.org/10.1016/j.paid.2018.06.003

Heinze, P. E., Fatfouta, R., & Schröder-Abé, M. (2020). Validation of an implicit measure of antagonistic narcissism. *Journal of Research in Personality*, *88*, 103993. https://doi.org/10.1016/j.jrp.2020.103993

Hernández-López, J. M., Santacreu-Mas, J., & Rubio-Franco, V. J. (1999). Evaluación de la personalidad: Una alternativa teórico-metodológica [Personality assessment: A theoretical-methodological alternative]. *Escritos de Psicología - Psychological Writings*, *1*(3), 20–28. https://doi.org/10.24310/espsiescpsi.vi3.13437

Hoaglin, D. C. (2016). Misunderstandings about Q and 'Cochran's Q test' in meta-analysis. *Statistics in Medicine*, *35*(4), 485–495. https://doi.org/10.1002/sim.6632

Hofstee, W. K. B. (1994). Who should own the definition of personality? *European Journal of Personality*, *8*(3), 149–162. https://doi.org/10.1002/per.2410080302

Holtzman, N. S. (2011). Facing a psychopath: Detecting the dark triad from emotionally-neutral faces, using prototypes from the Personality Faceaurus. *Journal of Research in Personality*, *45*(6), 648–654. https://doi.org/10.1016/j.jrp.2011.09.002

Holtzman, N. S., & Strube, M. J. (2013a). Above and beyond short-term mating, long-term mating is uniquely tied to human personality. *Evolutionary Psychology*, *11*(5), 1101–1129. https://doi.org/10.1177/147470491301100514

Holtzman, N. S., & Strube, M. J. (2013b). People with dark personalities tend to create a physically attractive veneer. *Social Psychological and Personality Science*, *4*(4), 461–467. https://doi.org/10.1177/1948550612461284

Hudek-Knezevic, J., Kardum, I., & Banov, K. (2023). The effects of the dark triad personality traits on health protective behaviours: Dyadic approach on self-reports and partner-reports. *Psychology & Health*, *38*(8), 987–1005. https://doi.org/10.1080/08870446.2021.1998497

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

Jones, D. N., & Paulhus, D. L. (2009). Machiavellianism. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 93–108). Guilford.

Jones, D. N., & Paulhus, D. L. (2011). Differentiating the dark triad within the interpersonal circumplex. In L. M. Horowitz & S. Strack (Eds.), *Handbook of interpersonal psychology: Theory, research, assessment, and therapeutic interventions* (pp. 249–268). Wiley.

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

Kandler, C., Bleidorn, W., Riemann, R., Spinath, F. M., Thiel, W., & Angleitner, A. (2010). Sources of cumulative continuity in personality: A longitudinal multiple-rater twin study. *Journal of Personality and Social Psychology*, *98*(6), 995–1008. https://doi.org/10.1037/a0019558

Kardum, I., Hudek-Knezevic, J., Marijanović, K., & Shackelford, T. K. (2023). Predicting mate poaching experiences from personality traits using a dyadic analysis. *The Journal of Sex Research*, *60*(3), 384–398. https://doi.org/10.1080/00224499.2022.2092586

Kardum, I., Hudek-Knezevic, J., & Mehic, N. (2022). Similarity indices of the dark triad traits based on self and partner-reports: Evidence from variable-centered and couple-centered approaches. *Personality and Individual Differences*, *193*, 111626. https://doi.org/10.1016/j.paid.2022.111626

Kardum, I., Hudek-Knezevic, J., Mehić, N., & Banov Trošelj, K. (2023). The dark triad traits and relationship satisfaction: Dyadic response surface analysis. *Journal of Personality*, *92*, 931–947. https://doi.org/10.1111/jopy.12857

Kay, C. S. (2021). The targets of all treachery: Delusional ideation, paranoia, and the need for uniqueness as mediators between two forms of narcissism and conspiracy belief. *Journal of Research in Personality*, *93*, 104128. https://doi.org/10.1016/j.jrp.2021.104128

Kholin, M., Meurs, J. A., Blickle, G., Wihler, A., Ewen, C., & Momm, T. D. (2016). Refining the openness–performance relationship: Construct specificity, contextualization, social skill, and the combination of trait self- and other-ratings. *Journal of Personality Assessment*, *98*(3), 277–288. https://doi.org/10.1080/00223891.2015.1076430

Kim, H., Di Domenico, S. I., & Connelly, B. S. (2019). Self-other agreement in personality reports: A meta-analytic comparison of self- and informant-report means. *Psychological Science*, *30*(1), 129–138. https://doi.org/10.1177/0956797618810000

Klipfel, K. M., & Kosson, D. S. (2018). The relationship between grandiosity, psychopathy, and narcissism in an offender sample. *International Journal of Offender Therapy and Comparative Criminology*, *62*(9), 2687–2708. https://doi.org/10.1177/0306624X17734784

Kyllonen, P. C., & Kell, H. (2018). Ability tests measure personality, personality tests measure ability: Disentangling construct and method in evaluating the relationship between personality and ability. *Journal of Intelligence*, *6*(3), 32. https://doi.org/10.3390/jintelligence6030032

Lämmle, L., Nussbeck, F. W., & Ziegler, M. (2021). Hello from the other side: Can we perceive others' darkness? Observers' accuracy of the dark triad. *Journal of Personality Assessment*, *103*(1), 106–119. https://doi.org/10.1080/00223891.2019.1683020

Larsen, R. J., Buss, D. M., Wismeijer, A., Song, J., & Van den Berg, S. (2017). *Personality psychology: Domains of knowledge about human nature*. McGraw Hill Education. https://research.utwente.nl/en/publications/personality-psychology-domains-of-knowledge-about-human-nature

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology*, *68*(1), 151–158. https://doi.org/10.1037/0022-3514.68.1.151

Lozano-Bleda, J. H. L., López Hernández, J. M., & Mas, J. S. (2010). Validez de constructo de pruebas objetivas de evaluación de la minusciosidad y la tendencia al riesgo. [Construct validity of consciousness and risk propensity objective tests.]. *EduPsykhé: Revista de Psicología y Educación*, *9*(2), 263–280.

Luan, Z., Poorthuis, A. M. G., Hutteman, R., Denissen, J. J. A., Asendorpf, J. B., & van Aken, M. A. G. (2019). Unique predictive power of other-rated personality: An 18-year longitudinal

study. *Journal of Personality*, *87*(3), 532–545. https://doi.org/10.1111/jopy.12413

Maaß, U., & Ziegler, M. (2017). Narcissistic self-promotion is not moderated by the strength of situational cues. *Personality and Individual Differences*, *104*, 482–488. https://doi.org/10.1016/j.paid.2016.09.008

Malesza, M., & Kaczmarek, M. C. (2020). The convergent validity between self- and peer-ratings of the dark triad personality. *Current Psychology*, *39*(6), 2166–2173. https://doi.org/10.1007/s12144-018-9906-7

McCrae, R. R. (2018). Method biases in single-source personality assessments. *Psychological Assessment*, *30*(9), 1160–1173. https://doi.org/10.1037/pas0000566

McDonald, J. D. (2008). Measuring personality constructs: The advantages and disadvantages of self- reports, informant reports and behavioural assessments. *Enquire*, *1*(1), 75–94. https://www.nottingham.ac.uk/sociology/documents/enquire/volume-1-issue-1-dodorico-mcdonald.pdf

Miller, J. D., Hyatt, C. S., Maples-Keller, J. L., Carter, N. T., & Lynam, D. R. (2017). Psychopathy and Machiavellianism: A distinction without a difference? *Journal of Personality*, *85*(4), 439–453. https://doi.org/10.1111/jopy.12251

Miller, J. D., Jones, S. E., & Lynam, D. R. (2011). Psychopathic traits from the perspective of self and informant reports: Is there evidence for a lack of insight? *Journal of Abnormal Psychology*, *120*(3), 758–764. https://doi.org/10.1037/a0022477

Mischel, W. (1968). *Personality and assessment.* Psychology Press. https://doi.org/10.4324/9780203763643

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, *8*(5), 336–341. https://doi.org/10.1016/j.ijsu.2010.02.007

Moor, L., & Anderson, J. R. (2019). A systematic literature review of the relationship between dark personality traits and antisocial online behaviours. *Personality and Individual Differences*, *144*, 40–55. https://doi.org/10.1016/j.paid.2019.02.027

Müller, S., & Moshagen, M. (2019). Controlling for response bias in self-ratings of personality: A comparison of impression management scales and the overclaiming technique. *Journal of Personality Assessment*, *101*(3), 229–236. https://doi.org/10.1080/00223891.2018.1451870

Muris, P., Meesters, C., & Timmermans, A. (2013). Some youths have a gloomy side: Correlates of the dark triad personality traits in non-clinical adolescents. *Child Psychiatry & Human Development*, *44*(5), 658–665. https://doi.org/10.1007/s10578-013-0359-9

Muris, P., Merckelbach, H., Otgaar, H., & Meijer, E. (2017). The malevolent side of human nature: A meta-analysis and critical review of the literature on the dark triad (narcissism, Machiavellianism, and psychopathy). *Perspectives on Psychological Science*, *12*(2), 183–204. https://doi.org/10.1177/1745691616666070

Nealis, L. J., Sherry, S. B., Lee-Baggley, D. L., Stewart, S. H., & Macneil, M. A. (2016). Revitalizing narcissistic perfectionism: Evidence of the reliability and the validity of an emerging construct. *Journal of Psychopathology and Behavioral Assessment*, *38*(3), 493–504. https://doi.org/10.1007/s10862-016-9537-y

Niemeyer, L. M., Grosz, M. P., Jallalvand, L., Mota, S., & Back, M. D. (2022). Toward a differentiated assessment of narcissism in forensic contexts: Validating the Narcissistic Admiration and Rivalry Questionnaire–Short Scale (NARQ-S) in a forensic sample. *Assessment*, *29*(4), 731–747. https://doi.org/10.1177/1073191120986608

Oltmanns, J. R., & Oltmanns, T. F. (2021). Self-other agreement on ratings of personality disorder symptoms and traits: Three meta-analyses. In T. D. Letzring & J. S. Spain (Eds.), *The Oxford handbook of accurate personality judgment* (pp. 276–293). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190912529.013.19

Oltmanns, J. R., & Widiger, T. A. (2021). The self-and informant-personality inventories for ICD-11: Agreement, structure, and relations with health, social, and satisfaction variables in older adults. *Psychological Assessment*, *33*(4), 300–310. https://doi.org/10.1037/pas0000982

O'Meara, A., Davies, J., & Hammond, S. (2011). The psychometric properties and utility of the Short Sadistic Impulse Scale (SSIS). *Psychological Assessment*, *23*(2), 523–531. https://doi.org/10.1037/a0022400

Ortner, T. M., & Proyer, R. T. (2015). Objective personality tests. In *Behavior-based assessment in psychology: Going beyond self-report in the personality, affective, motivation, and social domains* (pp. 133–149). Hogrefe. https://doi.org/10.1027/00437-000

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

Paulhus, D. L., Neumann, C. S., & Hare, R. D. (2009). *Manual for the Hare Self-Report Psychopathy Scale*. Multi-Health Systems.

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

Pfattheicher, S., Lazarević, L. B., Nielsen, Y. A., Westgate, E. C., Krstić, K., & Schindler, S. (2023). I enjoy hurting my classmates: On the relation of boredom and sadism in schools. *Journal of School Psychology*, *96*, 41–56. https://doi.org/10.1016/j.jsp.2022.10.008

Pineda, D., Galán, M., Martínez-Martínez, A., Campagne, D. M., & Piqueras, J. A. (2022). Same personality, new ways to abuse: How dark tetrad personalities are connected with cyber intimate partner violence. *Journal of Interpersonal Violence*, *37*(13–14), NP11223–NP11241. https://doi.org/10.1177/0886260521991307

Pineda, D., Martínez-Martínez, A., Galán, M., Rico-Bordera, P., & Piqueras, J. A. (2023). The dark tetrad and online sexual victimization: Enjoying in the distance. *Computers in Human Behavior*, *142*, 107659. https://doi.org/10.1016/j.chb.2023.107659

Pineda, D., Rico-Bordera, P., Martínez-Martínez, A., Galán, M., & Piqueras, J. A. (2022). Dark tetrad personality traits also play a role in bullying victimization. *Frontiers in Psychology*, *13*, 984744. https://doi.org/10.3389/fpsyg.2022.984744

Plouffe, R. A., Saklofske, D. H., & Smith, M. M. (2017). The assessment of sadistic personality: Preliminary psychometric evidence for a new measure. *Personality and Individual Differences*, *104*, 166–171. https://doi.org/10.1016/j.paid.2016.07.043

Prusik, M., & Szulawski, M. (2019). The Relationship Between the Dark Triad Personality Traits, Motivation at Work, and Burnout Among HR Recruitment Workers. *Frontiers Psychology*, *10*, 1290. https://doi.org/10.3389/fpsyg.2019.01290

Raskin, R. N., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports*, *45*(2), 590. https://doi.org/10.2466/pr0.1979.45.2.590

Rico-Bordera, P., Piqueras, J. A., Soto-Sanz, V., Rodríguez-Jiménez, T., Marzo, J. C., Galán, M., & Pineda, D. (2023). Civic engagement and personality: Associations with the big five and the dark triad. *International Journal of Environmental Research and Public Health*, *20*(3), 2126. https://doi.org/10.3390/ijerph20032126

Riemann, R., Angleitner, A., & Strelau, J. (1997). Genetic and environmental influences on personality: A study of twins reared together using the self- and peer-report NEO-FFI scales. *Journal of Personality*, *65*(3), 449–475. https://doi.org/10.1111/j.1467-6494.1997.tb00324.x

Rogier, G., Roberti, A., Garofalo, C., & Velotti, P. (2021). An investigation of spitefulness in violent offenders: Associations with the dark triad and emotion dysregulation. *Personality and Mental Health*, *15*(2), 89–99. https://doi.org/10.1002/pmh.1495

Roters, J., & Book, A. (2023). Using the HEXACO to explain the structure of borderline and psychopathic personality traits. *Personality and Mental Health*, *17*(4), 352–362. https://doi.org/10.1002/pmh.1584

Rubio, V. J., Santacreu, J., & Hernández, J.-M. (2004). The objective assessment of personality. An alternative to self-report based assessment. *Análisis y Modificación de Conducta*, *30*(134), 827–840. https://dialnet.unirioja.es/servlet/articulo?codigo=1149151

Sánchez-Bruno, A., & Borges del Rosal, Á. (2005). Transformación Z de Fisher para la determinación de intervalos de confianza del coeficiente de correlación de Pearson [Transformation of Fisher's Z to determine confidence intervals of Pearson's correlation coefficient]. *Psicothema*, *17*(1), 148–153.

Santacreu, J., & Hernández, J. M. (2018). Theory of personality: An interactive behavioral approach. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of personality and individual differences*. Springer.

Sherman, E. D., Miller, J. D., Few, L. R., Campbell, W. K., Widiger, T. A., Crego, C., & Lynam, D. R. (2015). Development of a short form of the Five-Factor Narcissism Inventory: The FFNI-SF. *Psychological Assessment*, *27*(3), 1110–1116. https://doi.org/10.1037/pas0000100

Spaans, M., Molendijk, M. L., de Beurs, E., Rinne, T., & Spinhoven, P. (2017). Self-reported personality traits in forensic populations: A meta-analysis. *Psychology, Crime & Law*, *23*(1), 56–78. https://doi.org/10.1080/1068316X.2016.1220555

Szabó, Z. P., Diller, S. J., Czibor, A., Restás, P., Jonas, E., & Frey, D. (2023). "One of these things is not like the others": The associations between dark triad personality traits, work attitudes, and work-related motivation. *Personality and Individual Differences*, *205*, 112098. https://doi.org/10.1016/j.paid.2023.112098

Thomas, L., & Egan, V. (2022). A systematic review and meta-analysis examining the relationship between everyday sadism and aggression: Can subclinical sadistic traits predict aggressive behaviour within the general population? *Aggression and Violent Behavior*, *65*, 101750. https://doi.org/10.1016/j.avb.2022.101750

Tyrer, P., Wang, D., Tyrer, H., Crawford, M., Loebenberg, G., Cooper, S., Barrett, B., & Sanatinia, R. (2021). Influence of apparently negative personality characteristics on the long-term outcome of health anxiety: Secondary analysis of a randomized controlled trial. *Personality and Mental Health*, *15*(1), 72–86. https://doi.org/10.1002/pmh.1496

Uzieblo, K., Decuyper, M., Bijttebier, P., & Verhofstadt, L. (2022). When the partner's reality bites: Associations between self- and partner ratings of psychopathic traits, relationship quality and conflict tactics. *International Journal of Offender Therapy and Comparative Criminology*, *66*(15), 1659–1681. https://doi.org/10.1177/0306624X221086560

Vandenbroucke, J. P., von Elm, E., Altman, D. G., Gøtzsche, P. C., Mulrow, C. D., Pocock, S. J., Poole, C., Schlesselman, J. J., Egger, M., & Initiative, for the S. (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. *Epidemiology*, *18*(6), 805–835. https://doi.org/10.1097/EDE.0b013e3181577511

Vander Molen, R. J., Kaplan, S., Choi, E., & Montoya, D. (2018). Judgments of the dark triad based on Facebook profiles. *Journal of Research in Personality*, *73*, 150–163. https://doi.org/10.1016/j.jrp.2017.11.010

Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality*, *40*(5), 472–481. https://doi.org/10.1016/j.jrp.2005.03.003

Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, *98*(2), 281–300. https://doi.org/10.1037/a0017908

Vazire, S., & Carlson, E. N. (2011). Others sometimes know us better than we know ourselves. *Current Directions in Psychological Science*, *20*(2), 104–108. https://doi.org/10.1177/0963721411402478

Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, *95*(5), 1202–1216. https://doi.org/10.1037/a0013314

Viechtbauer, W. (2010). Conducting meta-analyses in *R* with the metaphor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *The Lancet*, *370*(9596), 1453–1457. https://doi.org/10.1016/S0140-6736(07)61602-X

Walker, S. A., MacCann, C., & Jonason, P. K. (2023). The Dark Informant-Rated Triad (DIRT): A concise informant-rated measure of the dark triad. *European Journal of Psychological Assessment*. https://doi.org/10.1027/1015-5759/a000796

Webster, G., & Wongsomboon, V. (2020). The Hateful Eight (H8): An efficient multifaceted approach to the Short Dark Tetrad (SD4). PsyArXiv. https://doi.org/10.31234/osf.io/pr4u6

ASSESSING THE DARK PERSONALITY TRAITS WITH OBSERVER REPORTS: A META-ANALYSIS OF INTER-RATER AGREEMENT ON THE DARK TRIAD AND DARK TETRAD TRAITS

19

Weiss, B., Lavner, J. A., & Miller, J. D. (2018). Self- and partner-reported psychopathic traits' relations with couples' communication, marital satisfaction trajectories, and divorce in a longitudinal sample. *Personality Disorders*, 9(3), 239–249. https://doi.org/10.1037/per0000233

Weiss, B., Sleep, C. E., Lynam, D. R., & Miller, J. D. (2021). Evaluating the instantiation of narcissism components in contemporary measures of psychopathy. *Assessment*, 28(1), 15–28. https://doi.org/10.1177/1073191120916797