

Evaluating the Psychometric Properties of the Short Dark Triad (SD3)
in Italian Adults and Adolescents

Antonella Somma ^a, Delroy L. Paulhus ^b, Serena Borroni ^a, & Andrea Fossati ^a

a: Vita-Salute San Raffaele University, Milan, Italy, and San Raffaele Hospital, Milan, Italy

b: University of British Columbia, Vancouver, Canada

The final article is available at: Somma, A., Paulhus, D. L., Borroni, S., & Fossati, A. (2020). 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>

Correspondence concerning this article should be sent to Antonella Somma, Clinical Psychology and Psychotherapy Unit, IRCCS San Raffaele Turro Hospital, via Stamira d’Ancona 20, 20127, Milan, Italy. Telephone: +39 02 26433241; fax: +39 02 26433408; e-mail: somma.antonella@hsr.it

Summary

The term *Dark Triad* refers to three socially aversive personality dimensions (i.e., Machiavellianism, narcissism, and psychopathy) that are evident in the normal range of personality. Jones and Paulhus (2014) developed the Short Dark Triad (SD3) as a 27-item measure of the three constructs. To assess the psychometric properties of the Italian translation, 678 adult university students and 442 adolescent high school students were sampled. Cronbach α values for the subscales were acceptable in both samples. Subscale intercorrelations ranged from .29 to .55 in adults and .29 to .53 in adolescents. Although subscale means were higher in the adolescent sample, the two item correlation matrices did not differ significantly. A confirmatory factor analysis using multidimensional full-information item response theory showed that a three-correlated-factors model provided the best fit in both adults and adolescents. When controlled for overlap, SD3 subscales showed adequate convergent and discriminant validity coefficients in both samples. The current research contributes to the literature on dark personalities in two ways: (a) it provides detailed psychometric support for the Italian translation of SD3, and (b) it directly compares SD3 performance in younger and older students.

Keywords: Short Dark Triad; adults; adolescents; Item Response Theory.

Category = Multi-study Reports

Word count = 7,496

Introduction

Since its introduction by Paulhus and Williams (2002), the multi-dimensional construct labeled the *Dark Triad* has gained widespread popularity in social and personality psychology research. The term refers to three socially aversive personality dimensions that are evident in the normal range of personality (for a review, see Furnham, Richards, & Paulhus, 2013). Those three personality dimensions are: (a) *narcissism*, a trait characterized by manipulativeness, callousness, lack of empathy, impulsivity, and risk-taking; (b) *psychopathy*; a trait characterized by callous-unemotional traits, deceitfulness, impulsivity, and risk-taking; and (c) *Machiavellianism*, a trait characterized by strategic exploitation. Paulhus and Williams (2002) argued that, because they are intercorrelated, the three traits should be studied together. Otherwise, any apparent link with an outcome could be misattributed to the wrong predictor.

Although sound measures of narcissism (e.g., Narcissistic Personality Inventory; NPI, Raskin & Hall, 1979), psychopathy (e.g., Self Report Psychopathy scale; Paulhus, Neumann, & Hare, 2016), and Machiavellianism (e.g., Mach IV; Christie & Geis, 1970) already exist, the combined length of the available measures (124 items) was off-putting for many researchers; indeed, even with the shortest versions of each construct, the total number of items (i.e., 51) may still be taxing when time and space are at a premium (Jones & Paulhus, 2014).

Fortunately, shorter combination measures of all three constructs were soon developed. First to appear was the so-called "Dirty Dozen" questionnaire (DD; Jonason & Webster, 2010). Notwithstanding the adequate reliabilities reported for the three DD subscales, the DD showed some important limitations (Jones & Paulhus, 2014; Jonason et al., 2011; Rauthmann, 2013). For instance, the Machiavellianism subscale of the DD showed problematic discriminant validity with respect to psychopathy (e.g., Miller et al., 2012).

To overcome these limitations, Jones and Paulhus (2014) created the Short Dark Triad (SD3) as a 27-item measure of the three constructs. Based on a review of the relevant literature (Jones & Paulhus, 2011), a large item pool was assembled in order to ensure coverage of the key

aspects of each concept. Three theoretical principles motivated the item selection: (a) ego-identity goals drive narcissistic behavior, whereas instrumental goals drive Machiavellian and psychopathic behavior, (b) Machiavellianism differs from psychopathy with respect to temporal focus, and (c) all three have a callous core that encourages interpersonal manipulation. Item refinement and structural analyses finally reduced the original item set down to the final set of 27 items.

In a series of four studies that included a total of 1,063 community dwelling adults and college students, Jones and Paulhus (2014) reported adequate reliabilities (i.e., all Cronbach α values $> .70$) for the subscales. Both exploratory factor analysis and exploratory structural equation modeling revealed that three correlated latent factors could explain the observed correlations among the 27 SD3 items. In contrast to the DD, the SD3 subscales seemed to show adequate convergent and discriminant validity. A final validity study showed that SD3 self-reports converge with corresponding informant reports on all three subscales (Jones & Paulhus, 2014). Support for the SD3 has accumulated from a variety of other researchers (e.g., Book et al., 2015; Dowgwillo & Pincus, 2017; Gonclaves & Campbell, 2014; Schneider et al., 2017; Veselka et al., 2014).

Although translated into several other languages, the psychometric properties of the SD3 have not yet been evaluated in the Italian context. Moreover, the Dark Triad has rarely been studied in adolescents – regardless of the language. One exception is the study by Lau and Marsee (2013), which showed unique associations of delinquency with each of the triad components. Kerig and Stellwagen (2010) showed that associations with ‘theory of mind’ differed across the three components of the Dark Triad. Such research on younger samples can shed light on the different pathways leading to antisocial behavior, as well as offering preventive interventions (van Baardewijk, Vermeiren, Stegge, & Doreleijers, 2011). In the present study, we propose to fill those gaps in the literature.

Factor structure. Some investigators have questioned the distinctiveness of the Dark Triad – even using the SD3. Jonason and colleagues (2010) were among the first to contend that the three traits may best be viewed as facets of a single social orientation, namely, short-term exploitation of

others. In more recent publications, however, Jonason and colleagues have honored the distinctiveness of the constructs (Jonason et al., 2012). Others remain concerned about the factor structure of the SD3 (e.g., Miller et al., 2013).

Part of the concern about SD3 factor structure is shared with recent cautions regarding the application of latent variable modeling techniques to such domains as personality, psychopathology, and health (Reise & Waller, 2009). Nonetheless, IRT measurement models hold much promise for the development, psychometric analysis, refinement, and scoring of dysfunctional personality measures (Reise & Rodriguez, 2016). Based on considerations raised by Reise, Morizot, and Hays (2007) and Reise (2012), we relied on confirmatory full-information maximum likelihood IRT factor analysis to estimate the latent structure of SD3 items. In particular, we relied on Samejima's (1997) graded response model to fit the following three models: (a) a unidimensional model; (b) a three-factor model with orthogonal factors, in which SD3 items were forced to load on the corresponding factor they were a priori assigned to; and (c) a three-factor model with correlated factors, in which SD3 items were forced to load on the corresponding factor they were a priori assigned to. Moreover, the hypothesis of the measurement invariance across the adolescent and adult samples was formally tested in a multi-group IRT full-information maximum likelihood IRT factor analysis framework.

Nomological network. In order to extend the construct validity of the SD3 subscales, we evaluated their concurrent associations with several established personality constructs: (a) the traditional measure of Machiavellianism (Mach IV), (b) a measure of grandiose narcissism from the Five Factor Narcissism Inventory-Short Form (Sherman et al., 2015), and (c) a 3-factor measure of psychopathy, the Triarchic Psychopathy Measure (TriPM). The latter measure comprises *boldness*, *meanness*, and *disinhibition* (for a review, see Patrick and Drislane, 2015). Although never used in previous Dark Triad research, the latter two measures are sufficiently well established to use as sound criteria for assessing concurrent validity. Details of these measures may be found in the Method section below.

Predictions. Based on findings with the English version (Jones & Paulhus, 2014), the Italian SD3 Machiavellianism scale is expected to show a strong correlation with Mach IV; the SD3 narcissism scale is expected to show a strong positive correlation with FFNI-SF grandiose narcissism; and SD3 psychopathy is expected to show a strong positive correlation with the meanness and disinhibition facets of the TriPM.

Although no previous study has directly compared adolescent and adult samples, several studies have shown that, within samples, Dark Triad scores diminish with age (e.g., Jonason, Slomski, & Partyka, 2012; Vernon et al., 2008). Hence, we predict that the adolescent sample will score higher than the university sample on all three triad measures.

Materials and Methods

Participants

Sample 1. This sample was composed of 678 adult university students attending courses at a large State University in the Center of Italy. Participants responded to advertisements requesting potential volunteers for psychological studies taking place on campus and on the university web. Three hundred forty-seven (51.2%) participants were male and 331 (48.8%) participants were female; participant's mean age was 24.1 years, $SD = 2.64$ years (range: 19-30 years). In terms of participants' civil status, 670 (98.8%) were unmarried, and 8 (1.2%) were married.

Sample 2. This sample was composed of 442 adolescents attending large public high schools in Central Italy; 226 adolescents (51.1%) were female and 215 adolescents (48.6%) were male, whereas one adolescent (0.2%) did not report his/her gender. Adolescents' mean age was 16.06 years, $SD = 1.47$. Owing to space consideration, we relegated details of our missing data analysis to the Electronic Supplementary Material.

All participants gave their written consent to participate in the study after it had been explained to them; all participants were volunteers and received no financial or academic incentive to take part in the study.

Measures

Short Dark Triad (SD3; Jones & Paulhus, 2014). The SD3 is a 27-item, Likert-type, self-report measure that was specifically designed to assess three socially aversive personality traits, namely, narcissism (9 items), psychopathy (9 items), and Machiavellianism (9 items). Each item is measured on a five-point ordinal scale, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The SD3 yields separate scores for Machiavellianism, narcissism, and psychopathy scales. As noted earlier, its validity has been supported by structural analyses, concurrent correlations and informant ratings, and observable behavior (e.g., Jones & Paulhus, 2017).

Five Factor Narcissism Inventory-Short Form Grandiose Narcissism Scale (FFNI-SF GN; Sherman et al., 2015). The FFNI-SF is a 60-item, self-report measure of 15 traits related to grandiose and vulnerable narcissism: Acclaim-Seeking, Arrogance, Authoritativeness, Distrust, Entitlement, Exhibitionism, Exploitativeness, Grandiose Fantasies, Indifference, Lack of Empathy, Manipulativeness, Need for Admiration, Reactive Anger, Shame, and Thrill-Seeking. Previous studies (e.g., Sherman et al., 2015) showed that the FFNI-SF scales showed adequate internal consistency reliability estimates among undergraduates and clinical participants (i.e., all Cronbach α values $> .70$).

Grandiose narcissism is the sum of Acclaim-Seeking, Arrogance, Authoritativeness, Entitlement, Exhibitionism, Exploitativeness, and Grandiose Fantasies. *Vulnerable narcissism* is the sum of the remaining scales. Consistent with the 148-item FFNI (Glover et al., 2012), each FFNI-SF item is measured on a five-point ordinal scale ($1 = \text{Disagree strongly}$; $2 = \text{Disagree a little}$; $3 = \text{Neither agree nor disagree}$; $4 = \text{Agree a little}$; $5 = \text{Agree strongly}$). The Italian translation of the FFNI-SF showed good psychometric properties among University students (Fossati, Somma, Borroni, & Miller, 2017). Following Jones and Paulhus (2014), we relied on the items tapping into grandiose narcissism (FFNI-SF GN).

Triarchic Psychopathy Measure (TriPM; Patrick & Drislane, 2015). The TriPM is a 58-item self-report inventory divided into subscales indexing the three phenotypic components:

boldness, meanness, and disinhibition. Participants respond to each item on a 4-point Likert scale (3 = *true*, 2 = *mostly true*, 1 = *mostly false*, 0 = *false*). The Boldness scale is composed of 19 items that index tendencies toward social poise and effectiveness, emotional resiliency, and venturesomeness. The Disinhibition and Meanness scales (20 and 19 items, respectively) index broad disinhibition and callous-aggression factors, respectively. Published research provides support for the reliability (i.e., all Cronbach α values > .75) and validity of the TriPM as a measure of psychopathic features (e.g. Drislane, Patrick, & Arsal, 2014).

Of particular relevance here is that the TriPM has been validated in Italian samples (Sica et al., 2011; Somma, Borroni, Drislane, & Fossati, 2015). Given concerns about whether the Boldness factor is a genuine component of psychopathy (Evans & Tully, 2016; Neumann et al., 2013), we removed Boldness from the total score before using the TriPM as a criterion measure.

Machiavellianism (Mach IV; Christie & Geis, 1970). The Mach IV is a 20-item, Likert-type self-report measure that captures attitudes and behaviors associated with the Machiavellian personality construct. Each item is measured on a five-point scale, ranging from 1 (*Strongly disagree*) to 5 (*Strongly agree*). The instrument includes items tapping into Machiavellian tactics, cynicism, and morality. The Mach IV has typically yielded adequate internal consistency estimates of reliability (i.e., Cronbach α values > .75) and validation evidence has accumulated to include several hundred observer and behavioral studies (see Jones & Paulhus, 2009).

Owing to space consideration, we included a detailed description of the translation procedures and an extensive description of data analysis as Electronic Supplementary Material.

Results

Sample 1 and Sample 2 did not differ significantly in terms of gender proportion, $\chi^2(1) = 0.63, p > .40, \phi = .02$. Understandably, the samples significantly differed on age using a separate-variance $t(1094.99) = 64.93, p < .001$. Table 1 summarizes descriptive statistics, item-total correlations corrected for overlap (i.e., item convergent validity), and correlations of each item with the total score for the scale to which it was not assigned (i.e., item discriminant validity). On

average, SD3 items showed moderate, positive, albeit significant skewness values, *Median* = 0.45 (item 4), $p < .001$, min. (item 18) = -0.68, $p < .001$, max (item 25) = 1.49, $p < .001$. Four items -- namely, 2, 25, 26, and 27 - showed skewness values greater than 1.00. Median item kurtosis value was -0.61 (item 14), $p < .005$, min (item 23) = -1.08, $p < .001$; max (item 25) = 1.22.

Group differences. Table 2 summarizes descriptive statistics in the full sample and broken down by gender, Cronbach α coefficient values (as well as mean inter-item correlations), and scale inter-correlations for SD3 scales. Male participants and female participants did not add up to the total number of participants in Sample 2, because one Sample 2 participant refused to disclose his/her gender. A two-way MANOVA evidenced a significant omnibus difference between adult University students (i.e., Sample 1) and adolescent high school students (i.e., Sample 2), Pillai $V = .16$, $p < .001$, as well as between male and female participants, Pillai $V = .07$, $p < .001$, on SD3 scale average scores. There was no significant sample-by-gender interaction, Pillai $V = .00$, $p > .80$. As it can be observed in Table 2, male participants scored significantly higher than female participants in both samples, although the effect size estimates (i.e., Cohen d values) were in the small-to-moderate range (Cohen, 1988). According to Bonferroni multiple t -tests (Bonferroni-corrected nominal $p < .016$), adolescent high school students scored significantly higher than adult University students on SD3 Machiavellianism, $t(1118) = 10.84$, $p < .001$, $d = .65$, narcissism, $t(1118) = 5.28$, $p < .001$, $d = .32$, and psychopathy, $t(1118) = 12.86$, $p < .001$, $d = .77$ scales.

Box M test results suggested that the scale variance-covariance matrices (and hence the correlation matrices) did not differ by gender in either Sample 1, $M = 11.86$, $F(6, 3291142.95) = 1.97$, $p > .05$, or Sample 2, $M = 4.71$, $F(6, 1386975.92) = 0.78$, $p > .50$. Accordingly, we pooled males and females before calculating the correlation matrices in both Sample 1 and Sample 2.

Finally, Box M test results suggested that the item variance-covariance matrix computed in adults was significantly different from that computed in the adolescents, $M = 20.89$, $F(6, 5976192.23) = 3.47$, $p < .01$. However, when scale scores were standardized within each sample (i.e., they resulted in two correlation matrices), the Box M test for the comparison between Sample

1 and Sample 2 on scale covariance (actually, correlation) matrix became non-significant, $M = 11.16$, $F(6, 5976192.23) = 1.86$, $p > .05$. Indeed, when we formally compared bivariate correlations among subscales that were computed in Sample 1 and in Sample 2, respectively, none of the Fisher z tests reached statistical significance -- even with no Bonferroni correction of the nominal p -value (i.e., $p < .05$), min. Fisher z value = 0.35 (correlation between SD3 narcissism and psychopathy scales), max. Fisher z value = 1.94 (correlation between SD3 Machiavellianism and psychopathy scales), all $ps > .05$.

In the present study, full-information maximum likelihood IRT factor analysis models and indices were estimated using the ‘mirt’ R package (Chalmers et al., 2012) whereas rest-score functions were computed using the ‘mokken’ R package (van der Ark, 2014). Goodness-of-fit indices of multidimensional confirmatory full-information IRT factor models of SD3 items are summarized in Table 3. Rest-score functions (and confidence bands) for all items in Sample 1 and Sample 2 are displayed in Figure 1 and Figure 2, respectively. Item discrimination and threshold estimates based on multidimensional confirmatory full-information IRT three correlated factor model in Sample 1 and in Sample 2 are listed in Table 4.

Finally, in order to test the invariance of SD3 IRT factor structure (and detect differential item functioning) across the adolescent and adult samples, we performed IRT multi-group full-information maximum likelihood factor analysis (Chalmers, 2012). First, we fitted a “configural” invariance model (i.e., a model postulating the invariance of the number of factors across groups). This model showed adequate values of fit indices, $M_2^*(480) = 1647.19$, $RMSEA_2 = .047$, 90% CI $RMSEA_2 = 0.44, 0.49$, $BIC = 84017.01$. Then, we turned to a scalar invariance model to test for measurement invariance¹, that is, a model with invariant factor loadings and thresholds across samples. We observed adequate values of fit statistics even for the most restrictive model, $M_2^*(612) = 2307.92$, $RMSEA_2 = .049$, 90% CI $RMSEA_2 = 0.47, 0.51$, $BIC = 83766.79$, supporting the

¹ SD3 items factor loadings and thresholds were set as invariant in tandem given that both parameters affect the item characteristic curve simultaneously (e.g., Millsap & Yun-Tein, 2004).

hypothesis of measurement invariance of the IRT three-factor structure across subgroups based on participants' age.

Unique components. Consistent with previous research on the English version, the SD3 subscales showed substantial positive intercorrelations (see Table 2; median $r = .49$ and $.42$ in adults and adolescents, respectively). To isolate their unique components, residualized measures of narcissism, Machiavellianism, and psychopathy were created. Descriptive statistics, internal consistency estimates (i.e., Cronbach α values), for the Mach IV total score, FFNI-SF GN and TriPM total score in Sample 1 and in Sample 2 are listed in Table 5. Correlations (Pearson r coefficients) between these residualized SD3 scale scores and their corresponding criteria are also listed in Table 5. In each sample, the nominal significance level (i.e., $p < .05$) was corrected according to the Bonferroni procedure and set at $p < .0016$. Bold highlights indicate convergent validity coefficients. Within each row, different superscripts indicate that convergent validity coefficients were significantly different from discriminant validity coefficients according to Steiger's z test.

Disattenuating for measurement error better highlights these convergent values. In Sample 1, the convergent validity coefficients become $.72$, $.74$, and $.40$ for Machiavellianism, narcissism, and psychopathy scales. In Sample 2, the disattenuated values become $.68$, $.89$, and $.73$ for Machiavellianism, narcissism, and psychopathy, respectively.

Note that our comparison of convergent and discriminant validities was hampered to some degree by the overlap of the three criterion variables. In Sample 1, the Mach IV total score correlated $.33$ and $.23$, all $ps < .001$, with FFNI-SF GN and TriPM total score, respectively, whereas FFNI-SF GN score correlated $.58$, $p < .001$, with TriPM total score. Among adolescent high school students (i.e., Sample 2 participants), the Mach IV total score correlated $.37$ and $.50$, all $ps < .001$, with FFNI-SF GN and TriPM total score, respectively; the correlation of FFNI-SF GN score with TriPM total score was $.58$, $p < .001$.

Discussion

The current study contributes to the literature on dark personalities in two ways. First, it provides support for the Italian translation of the Short Dark Triad (SD3) personality measure. As with the English version, our findings suggested that the Italian items tap three latent dimensions of malevolent personality corresponding to narcissism, Machiavellianism, and psychopathy.

Second, ours is the first study to compare the performance of the same Dark Triad instrument in adolescent and adult participants. Although both groups were students, the younger group (mean age 16) scored significantly higher than the older group (mean age 24) on all three subscales. Internal consistency reliability estimates (i.e., Cronbach α values) were slightly lower among adolescents (.67-.69) than among adult university students (.77-.84). Most important, however, the two variance-covariance matrices did not differ significantly. Hence, the overall pattern of item relationships, while not as sharply defined as for the English SD3 (i.e., Jones & Paulhus, 2014), was similar for adults and adolescents.

In the adult sample, item-total correlations corrected for item-total overlap (i.e., item convergent validity coefficients) were significant for all items in our adult University student sample (i.e., Sample 1). However, not all findings were in accordance with those obtained for the English version. For example, 12 of the convergent validity coefficients failed to exceed the corresponding discriminant validity coefficients (i.e., correlations with a subscale to which it was not assigned). Nonetheless, the mean convergent correlation (.44) exceeded the mean discriminant correlation (.27). Similar considerations held for our adolescent sample in which the majority of item convergent validity coefficients failed to exceed the corresponding discriminant validity coefficients. Nonetheless, the mean convergent correlation (.32) was higher than the mean discriminant correlation (.19). On the whole, then, the Italian SD3 items appear to capture the dimension to which they were assigned.

Because the three SD3 factors were designed to be oblique, cross loadings are expected to appear in replication samples – even in the same language. In the case of an item set translated into

Italian, it is difficult to say whether the higher rate of cross loadings indicates (a) differences in culture, (b) differences in language connotation, or (c) just a matter of chance.

Factor structure. Consistent with Jones and Paulhus (2014), our multidimensional IRT confirmatory factor analysis findings suggested the adequacy of the SD3 three-factor structure. Indeed, all fit indices suggested retention of the three correlated factors model as the best fitting model in both adults and adolescents. Although the three subscales represented dissociable latent dimensions, the corresponding IRT factors were substantially inter-correlated in both samples. Although there is consensus that the overlap in the English version results from a common core, interpretations of that core range from callousness (Jones & Figueredo, 2015), honesty-humility (Visser et al., 2016), antagonism (Jonason & Tost, 2010), or simply psychopathy (Glenn & Sellbom, 2015).

In our adult sample (see Table 4), only 7 of 27 SD3 items failed to show substantial item discrimination parameters (i.e., a coefficient values of .90 or greater; Sharp et al., 2014) on their corresponding factors. That number rose to 12 in our adolescent sample. This finding seemed to suggest that the Italian SD3 may be more appropriate for use with adult subjects. Notably, the IRT three-factor structure showed measurement invariance across adolescent and adult samples (thus, making mean comparisons legitimate), at least according to the results of our multiple-group IRT analysis.

Nomological network. Overall, results of convergent-discriminant validation supported the differentiation of the SD3 subscales. Extending previous evidence on convergent validity of the English subscales (Jones & Paulhus, 2014; Miller et al., 2015), our data suggested that (even the) residualized Machiavellianism scale provides at least moderate convergence with the gold standard measure of Machiavellianism (Mach IV) in both adult University students and adolescent high school students. The disattenuated value was even stronger.

Confirming and extending evidence of convergence between SD3 narcissism and the NPI (Jones & Paulhus, 2014), we found that the FFNI-SF GN scale scores showed a robust association

with the residualized narcissism scale score. Moreover, the SD3 narcissism scale provided adequate discriminant validity with respect to self-reports of both Machiavellian personality features (i.e., Mach IV total score) and psychopathy traits (i.e., TriPM total score). Similar results held for the two samples. Finally, the residualized psychopathy scale showed adequate convergent validity with respect to the TriPM, and satisfactory discriminant validity with respect to both FFNI-SF GN and Mach IV total scores in both adults and adolescents. Again, disattenuated values were even stronger.

In summary, although we relied on criterion measures of narcissism and psychopathy that differed from those used to validate the English version (Jones & Paulhus, 2014), we feel that our findings on the convergent-discriminant validity provided adequate support for the Italian version. Our criterion measures -- Mach IV, FFNI-SF GN scale, and TriPM -- are all well-established measures. Any difficulties we found with discriminant validities may be attributable, in part, to substantial overlap observed among the criterion measures: Although a replication study is in order, we see utility in the Italian SD3 for research as well as for screening purposes. As with the English version, the measure facilitates quick and valid assessment when time and space are limited.

Limitations. Of course, our findings should be considered in the light of several limitations. Although it involved a large number of participants ($N = 1,120$), our research was based on volunteers. Both adult and adolescent samples were convenient study groups rather than samples representative of the Italian population. Moreover, we relied only on student participants in both the adult and adolescent samples. This sampling procedure could potentially limit the generalizability of our findings. Completely unknown is the degree to which the findings can be extended to clinical and forensic populations.

Finally, we relied only on a limited number of criterion measures in order to evaluate convergent-discriminant validity. Although the FFNI, Mach IV, and TriPM are well-established measures of narcissism, Machiavellianism, and psychopathy, respectively, they all have their critics. Debates continue over the appropriate constituents of narcissism (e.g., grandiose vs. vulnerable

aspects; Miller, Lynam, Hyatt, & Campbell, 2017) and psychopathy (e.g., the relevance of boldness; Carré, Mueller, Schleicher, & Jones, 2018; Lilienfeld et al., 2016). Hence, future research should include alternative measures of narcissism (e.g., the NPI to measure narcissism, and the Self Report Psychopathy scale [Paulhus et al., 2016] to capture psychopathy).

Moreover, our criterion measures were all self-report instruments. Hence, we were not able to disentangle the effect of shared construct variance from the effect of shared method variance (Nunnally & Bernstein, 1994). Accordingly, our convergent-discriminant validity data need be replicated using alternative modes of measurement, for example, peer ratings (Jones & Paulhus, 2014) and overt behavior (Jones & Paulhus, 2017).

References

- Book, A., Visser, B. A., & Volk, A. A. (2015). Unpacking “evil”: Claiming the core of the Dark Triad. *Personality and Individual Differences*, 73, 29-38. doi: 10.1016/j.paid.2014.09.016
- Carré, J. R., Mueller, S. M., Schleicher, K. M., & Jones, D. N. (2018). Psychopathy and deviant workplace behavior: A comparison of two psychopathy models. *Journal of Personality Disorders*, 32, 242-261. doi: 10.1521/pedi_2017_31_296
- Chalmers, R.P. (2012). mirt: A Multidimensional Item Response Theory package for the R environment. *Journal of Statistical Software*, 48, 1-29. doi:10.18637/jss.v048.i06
- Chen, W. H., & Thissen, D. (1997). Local dependence indices for item pairs using item response theory. *Journal of Educational and Behavioral Statistics*, 22, 265-289. doi:10.3102/10769986022003265
- Christie, R., & Geis, F. (1970). *Studies in Machiavellianism*. New York, NY: Academic Press.
- Dowgwillo, E.A., & Pincus, A.L. (2017). Differentiating dark triad traits within and across interpersonal circumplex surfaces. *Assessment*, 24, 24-44. doi:10.1177/1073191116643161
- Drislane, L.E., Patrick, C.J., & Arsal, G. (2014). Clarifying the content coverage of differing psychopathy inventories through reference to the Triarchic Psychopathy Measure. *Psychological Assessment*, 26, 350-362. doi:10.1037/a0035152
- Evans, L., & Tully, R. J. (2016). The Triarchic Psychopathy Measure (TriPM): Alternative to the PCL-R? *Aggression and Violent Behavior*, 27, 79-86. doi:10.1016/j.avb.2016.03.004
- Fossati, A., Pincus, A.L., Borroni, S., Munteanu, A., & Maffei, C. (2014). Are pathological narcissism and psychopathy different constructs or different names for the same thing? A study based on Italian nonclinical adult participants. *Journal of Personality Disorders*, 28, 394–418. doi:10.1521/pedi_2014_28_127
- Furnham, A., Richards, S.C, & Paulhus, D.L. (2013). The Dark Triad of personality: A 10-year review. *Social and Personality Psychology Compass*, 7, 199-216. doi:10.1111/spc3.12018

- Glover, N., Miller, J. D., Lynam, D. R., Crego, C., & Widiger, T. A. (2012). The five-factor narcissism inventory: A five-factor measure of narcissistic personality traits. *Journal of Personality Assessment*, 94, 500-512. doi:10.1080/00223891.2012.670680
- Jonason, P. K., Slomski, S., & Partyka, J. (2012). The Dark Triad at work: How toxic employees get their way. *Personality and Individual Differences*, 52, 449-453. doi:10.1016/j.paid.2011.11.008
- Jonason, P. K., & Tost, J. (2010). I just cannot control myself: The Dark Triad and self-control. *Personality and Individual Differences*, 49, 611–615. doi:10.1016/j.paid.2010.05.031
- Jonason, P. K., & Webster, G. D. (2010). The Dirty Dozen: A concise measure of the Dark Triad. *Psychological Assessment*, 22, 420-432. doi: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). New York: 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). New York: Wiley.
- Jones, D.N., & Paulhus, D.L. (2014). Introducing the short dark triad (SD3): A brief measure of dark personality traits. *Assessment*, 21, 28-41. doi:10.1177/1073191113514105
- Jones, D.N., & Paulhus, D.L. (2017). Duplicity among the Dark Triad: Three faces of deceit. *Journal of Personality and Social Psychology*, 113, 329-342. doi:10.1037/pspp0000139
- Kerig, P.K., & Stellwagen, K.K. (2010). Roles of callous-unemotional traits, narcissism, and Machiavellianism in childhood aggression. *Journal of Psychopathology and Behavioral Assessment*, 32, 343-352. doi: 10.1007/s10862-009-9168-7
- Lau, K.S.L., & Marsee, M.A. (2013). Exploring narcissism, Machiavellianism, and psychopathy in youth: Examination of associations with antisocial behavior and aggression. *Journal of Child and Family Studies*, 22, 355–367. doi:10.1007/s10826-012-9586-0

- Lilienfeld, S.O., Smith, S.F., Sauvigné, K.C., Patrick, C.J., Drislane, L.E., Latzman, R.D., & Krueger, R.F. (2016). Is boldness relevant to psychopathic personality? Meta-analytic relations with non-Psychopathy Checklist-based measures of psychopathy. *Psychological Assessment, 28*, 1172-1185. doi:10.1037/pas0000244
- McDonald, R.P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.
- Miller, J.D., Few, L.R., Seibert, L.A., Watts, A., Zeichner, A., & Lynam, D. R. (2012). An examination of the Dirty Dozen measure of psychopathy: A cautionary tale about the costs of brief measures. *Psychological Assessment, 24*, 1048-1053. doi:10.1037/a0028583
- 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*, 439-453.
- Millsap, R.E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research, 39*, 479-515.
- Neumann, C.S., Uzieblo, K., Crombez, G., & Hare, R. D. (2013). Understanding the Psychopathic Personality Inventory (PPI) in terms of the unidimensionality, orthogonality, and construct validity of PPI-I and -II. *Personality Disorders: Theory, Research, and Treatment, 4*, 77–79. doi:10.1037/a0027196
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Patrick, C. J., & Drislane, L. E. (2015). Triarchic model of psychopathy: Origins, operationalizations, and observed linkages with personality and general psychopathology. *Journal of Personality, 83*, 627-643. doi:10.1111/jopy.12119
- Paulhus, D. L., & Williams, K. M. (2002). The Dark Triad of personality: narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality, 36*, 556-563. doi:10.1016/S0092-6566(02)00505-6
- Paulhus, D. L., Neumann, C. S., & Hare, R. D. (2016). *Manual for the Self-Report Psychopathy scale* (4th edition). Toronto: Multi-Health Systems.

- Pincus, A. L. (2013). The Pathological Narcissism Inventory. In J. Ogrudniczuk (Ed.), *Understanding and treating pathological narcissism* (pp. 93–110). Washington, DC: American Psychological Association.
- Pincus, A. L., & Lukowitsky, M. R. (2010). Pathological narcissism and narcissistic personality disorder. *Annual Review of Clinical Psychology*, 6, 421–446.
doi:10.1146/annurev.clinpsy.121208.131215
- Raskin, R., & Hall, C. S. (1979). A Narcissistic Personality Inventory. *Psychological Reports*, 45, 590. doi:10.2466/pr0.1979.45.2.590
- Reise, S. P., & Rodriguez, A. (2016). Item response theory and the measurement of psychiatric constructs: some empirical and conceptual issues and challenges. *Psychological Medicine*, 46; 2025-2039. doi:10.1017/S0033291716000520
- Reise, S. P., & Waller, N. G. (2009). Item response theory and clinical measurement. *Annual Review of Clinical Psychology*, 5, 27-48. doi:10.1146/annurev.clinpsy.032408.153553.
- Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, 95, 129-140. doi:10.1080/00223891.2012.725437
- Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, 92, 544–559. doi:10.1080/00223891.2010.496477
- Salekin, R. T., Worley, C., & Grimes, R. D. (2010). Treatment of psychopathy: A review and brief introduction to the mental model approach for psychopathy. *Behavioral Sciences & The Law*, 28, 235-266. doi:10.1002/bsl.928
- Samejima, F. (1997). Graded response model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of item response theory* (pp. 85-100). New York, NY: Springer-Verlag.
- Schneider, T. J., McLarnon, M. J., & Carswell, J. J. (2017). Career Interests, Personality, and the Dark Triad. *Journal of Career Assessment*, 25, 338-351. doi:10.1177/1069072715616128

- 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, 1110-1116. doi:10.1037/pas0000100
- Sica, C., Drislane, L., Caudek, C., Angrilli, A., Bottesi, G., Cerea, S., & Ghisi, M. (2015). A test of the construct validity of the Triarchic Psychopathy Measure in an Italian community sample. *Personality and Individual Differences*, 82, 163-168. doi:10.1016/j.paid.2015.03.015
- Somma, A., Borroni, S., Drislane, L.E., & Fossati, A. (2016). Assessing the triarchic model of psychopathy in adolescence: Reliability and validity of the Triarchic Psychopathy Measure (TriPM) in three samples of Italian community-dwelling adolescents. *Psychological Assessment*, 28, e36-e43. doi:10.1037/pas0000184
- van Baardewijk, Y., Vermeiren, R., Stegge, H., & Doreleijers, T. (2011). Self-reported psychopathic traits in children: Their stability and concurrent and prospective association with conduct problems and aggression. *Journal of Psychopathology and Behavioral Assessment*, 33, 236-245. doi:10.1007/s10862-010-9215-4
- van der Ark, L.A. (2014). New developments in Mokken scale analysis in R. *Journal of Statistical Software* 48, 1–27.
- Vernon, P. A., Villani, V. C., Vickers, L. C., & Harris, J. A. (2008). A behavioral genetic investigation of the Dark Triad and the Big 5. *Personality and Individual Differences*, 44, 445-452. doi:10.1016/j.paid.2007.09.007
- Veselka, L., Giammarco, E. A., & Vernon, P. A. (2014). The Dark Triad and the seven deadly sins. *Personality and Individual Differences*, 67, 75-80. doi:10.1016/j.paid.2014.01.055

Table 1

SD3 Item Analyses in Adult and Adolescent Samples: Descriptive Statistics and Part-Whole Correlations.

Short Dark Triad Scale Items	Sample 1 (N=678 Adult University Students)					Sample 2 (N=442 Adolescent Students)				
	<i>M</i>	<i>SD</i>	Mach	Narc	Psych	<i>M</i>	<i>SD</i>	Mach	Narc	Psych
1. It's not wise...(Machiavellianism)	3.40	1.06	.26* ^a	-.04	.06	3.34	1.16	.17 ^a	-.05	.05
2. I like to use clever...(Machiavellianism)	1.85	1.04	.54* ^a	.30 ^a	.39 ^a	2.33	1.23	.41 ^a	.35 ^a	.34 ^a
3. Whatever it takes...(Machiavellianism)	2.03	1.03	.59* ^a	.34 ^a	.41 ^a	2.80	1.19	.40* ^a	.18 ^a	.25 ^a
4. Avoid direct conflict...(Machiavellianism)	2.43	1.15	.52* ^a	.20 ^a	.25 ^a	2.88	1.10	.21* ^a	-.01	-.05
5. It's wise to keep track...(Machiavellianism)	2.37	1.12	.65* ^a	.36 ^a	.43 ^a	2.94	1.17	.42* ^a	.16	.30 ^a
6. You should wait...(Machiavellianism)	2.40	1.23	.56 ^a	.27 ^a	.51 ^a	3.14	1.28	.40 ^a	.19 ^a	.36 ^a
7. There are things...(Machiavellianism)	3.56	1.06	.47* ^a	.13	.25 ^a	3.74	1.01	.29* ^a	.03	.17 ^a
8. Make sure your plans...(Machiavellianism)	2.58	1.11	.57* ^a	.32 ^a	.42 ^a	3.22	1.15	.38* ^a	.14	.24 ^a
9. Most people...(Machiavellianism)	2.83	1.21	.49* ^a	.36 ^a	.40 ^a	3.02	1.20	.33 ^a	.27 ^a	.27 ^a
10. People see me...(Narcissism)	2.42	0.99	.21 ^a	.45* ^a	.24 ^a	2.44	1.10	.19 ^a	.49* ^a	.34 ^a
11. I hate being the center...(R; Narcissism)	2.97	1.12	.04	.32* ^a	.18 ^a	2.80	1.15	.02	.18 ^a	.15
12. Many group activities...(Narcissism)	2.12	0.90	.27 ^a	.38 ^a	.33 ^a	2.67	1.07	.18 ^a	.40 ^a	.31 ^a
13. I know that I am special...(Narcissism)	2.12	0.91	.22 ^a	.45* ^a	.27 ^a	2.55	1.13	.07	.39* ^a	.26 ^a
14. I like to get acquainted...(Narcissism)	2.93	1.09	.43 ^a	.40 ^a	.31 ^a	3.22	1.11	.30 ^a	.26 ^a	.28 ^a
15. I feel embarrassed...(R; Narcissism)	2.77	1.15	.05	.27* ^a	.15 ^a	2.57	1.19	.03	.24 ^a	.18 ^a

16. I have been compared to... (Narcissism)	2.38	1.17	.28 ^a	.37^a	.38 ^a	2.90	1.25	.17 ^a	.36^{*a}	.22 ^a
17. I am an average person (R; Narcissism)	2.54	1.14	.16 ^a	.35^{*a}	.23 ^a	2.38	1.29	.13	.37^{*a}	.19 ^a
18. I insist on getting...(Narcissism)	3.46	1.04	.22 ^a	.19^a	.25 ^a	3.80	1.02	.21 ^a	.09	.17 ^a
19. I like to get revenge...(Psychopathy)	2.55	1.18	.42 ^a	.43 ^a	.49^a	2.88	1.19	.33 ^a	.34 ^a	.43^a
20. I avoid dangerous situations (R; Psychopathy)	2.65	1.21	-.05	.18 ^a	.14^a	2.85	1.19	.12	.18 ^a	.27^a
21. Payback needs to be quick...(Psychopathy)	2.01	1.03	.28 ^a	.22 ^a	.38^{*a}	2.64	1.27	.26 ^a	.19 ^a	.35^a
22. People often say...(Psychopathy)	2.09	1.11	.35 ^a	.17 ^a	.48^{*a}	2.79	1.26	.16	.08	.30^{*a}
23. It's true that I can be mean...(Psychopathy)	2.80	1.23	.48 ^a	.21 ^a	.43^a	3.13	1.19	.29 ^a	.20 ^a	.33^a
24. People who mess with me...(Psychopathy)	2.46	1.15	.43 ^a	.40 ^a	.49^a	2.94	1.15	.29 ^a	.26 ^a	.39^a
25. I have never gotten...(R; Psychopathy)	1.78	1.16	.06	.19 ^a	.27^a	2.18	1.40	.02	.14	.13
26. I enjoy having sex...(Psychopathy)	1.85	1.14	.30 ^a	.33 ^a	.38^a	2.29	1.39	.18 ^a	.33 ^a	.26^a
27. I'll say anything to get what I...(Psychopathy)	1.83	0.95	.48 ^a	.29 ^a	.42^a	2.74	1.29	.27 ^a	.36 ^a	.34^a

Note. Bold indicates item-total correlations corrected for item-total overlap. Correlations coefficients equal to, or greater than .13 and .16 in absolute value in Sample 1 and Sample 2, respectively are significant at Bonferroni-corrected p -level (i.e., $p < .0006$). Critical values for a relevant difference between convergent validity (i.e., r_{i-i}) coefficients and discriminant validity coefficients were .08 and .10 in Sample 1 and Sample 2, respectively. Asterisks indicate items whose r_{i-i} values were significantly greater than all discriminant validity coefficients. a: $p < .0006$.

Table 2

SD3 Descriptive Statistics by Gender in Adult and Adolescent Samples.

Sample 1 (Adult University Students)												
	Whole Sample (<i>N</i> = 678)			Male Participants (<i>n</i> = 347)			Female Participants (<i>n</i> = 331)			ES	Correlations	
Short Dark Triad Scales	<i>M</i>	<i>SD</i>	α (MIC)	<i>M</i>	<i>SD</i>	α (MIC)	<i>M</i>	<i>SD</i>	α (MIC)	<i>d</i>	<i>1</i>	<i>2</i>
1. Machiavellianism	2.60	0.71	.84 (.38)	2.70	0.73	.84 (.37)	2.50	0.67	.83 (.34)	0.29***	--	
2. Narcissism	2.63	0.56	.87 (.43)	2.76	0.51	.87 (.42)	2.50	0.57	.87 (.44)	0.47***	.39***	--
3. Psychopathy	2.23	0.62	.77 (.28)	2.34	0.60	.73 (.23)	2.10	0.62	.80 (.31)	0.39***	.55***	.49***
Sample 2 ¹ (Adolescent High School Students)												
	Pooled Sample (<i>N</i> = 442)			Male Participants (<i>n</i> = 215)			Female Participants (<i>n</i> = 226)			ES	Correlations	
Short Dark Triad Scales	<i>M</i>	<i>SD</i>	α (MIC)	<i>M</i>	<i>SD</i>	α (MIC)	<i>M</i>	<i>SD</i>	α (MIC)	<i>d</i>	<i>1</i>	<i>2</i>
1. Machiavellianism	3.05	0.60	.69 (.20)	3.14	0.61	.68 (.19)	2.96	0.59	.69 (.20)	0.31***	.53	
2. Narcissism	2.82	0.57	.66 (.18)	2.96	0.57	.63 (.16)	2.67	0.54	.65 (.17)	0.53***	.29***	.56
3. Psychopathy	2.72	0.63	.67 (.19)	2.86	0.62	.63 (.16)	2.58	0.61	.70 (.20)	0.47***	.42***	.47***

Note. ES: effect size estimate; male participants and female participants do not add up to the total participants in Sample 2, because one participant did not disclose his/her gender.

*** $p < .001$.

Table 3.
Multidimensional IRT Confirmatory Factor Analyses of the SD3 Items in Adult and Adolescent Samples.

Sample 1 ($N=678$, Adult University Students)								
SD3 Item Models	M_2^*	df	RMSEA ₂	90% CI RMSEA ₂	SRMSR	AIC	BIC	SABIC
a) 1 factor	1566.70***	243	.090	.085 - .094	.084	48226.91	48837.00	48408.36
b) 3 uncorrelated factors	1577.31***	243	.090	.086 - .094	.168	48462.49	49072.57	48643.93
c) 3 correlated factors	1103.04***	240	.073	.068 - .077	.076	47877.95	48501.59	48063.42
Sample 2 ($N=442$, Adolescent High School Students)								
SD3 Item Models	M_2^*	df	RMSEA ₂	90% CI RMSEA ₂	SRMSR	AIC	BIC	SABIC
a) 1 factor	596.90***	243	.057	.052 - .063	.074	34929.96	35482.29	35053.86
b) 3 uncorrelated factors	519.73***	243	.051	.045 - .057	.117	35001.25	35553.58	35125.15
c) 3 correlated factors	443.03***	240	.043	.037 - .050	.069	34761.09	35325.70	34887.75

Note. RMSEA₂: sample bivariate root mean square error of approximation; CI: confidence interval; SRMSR: standardized root mean square residual; AIC: Akaike information criterion; BIC: Bayesian information criterion; SABIC: sample-size adjusted Bayesian information criterion.

*** $p < .001$

Table 4.

Multidimensional IRT Confirmatory Factor Analyses: Item Discrimination and Item Threshold Estimates in University and High School Students.

Short Dark Triad Scale Items	Sample 1 (N=678, Adult University Students)							Sample 2 (N=442, Adolescent High School Students)						
	Mach	Narc	Psych					Mach	Narc	Psych				
	<i>a1</i>	<i>a2</i>	<i>a3</i>	<i>d1</i>	<i>d2</i>	<i>d3</i>	<i>d4</i>	<i>a1</i>	<i>a2</i>	<i>a3</i>	<i>d1</i>	<i>d2</i>	<i>d3</i>	<i>d4</i>
1. It's not wise...(Machiavellianism)	0.42***	--	--	-1.75	-0.10	1.36	3.32	0.30*	--	--	-1.55	-0.15	1.23	2.51
2. I like to use clever...(Machiavellianism)	1.62***	--	--	-5.09	-3.31	-1.96	-0.03	1.21***	--	--	-3.42	-1.81	-0.64	0.90
3. Whatever it takes...(Machiavellianism)	1.72***	--	--	-4.97	-3.58	-1.52	0.65	0.92***	--	--	-2.71	-1.11	0.48	1.85
4. Avoid direct...(Machiavellianism)	1.31***	--	--	-3.78	-1.95	-0.42	1.40	0.32**	--	--	-2.75	-0.87	0.59	2.03
5. It's wise to keep...(Machiavellianism)	2.26***	--	--	-5.34	-2.86	-0.75	1.64	1.17***	--	--	-2.83	-0.94	0.76	2.23
6. You should wait...(Machiavellianism)	1.78***	--	--	-3.76	-2.32	-0.62	1.12	1.41***	--	--	-2.12	-0.49	0.87	2.56
7. There are things...(Machiavellianism)	1.03***	--	--	-1.89	0.33	1.89	3.40	0.68***	--	--	-1.29	0.62	2.14	3.78
8. Make sure your...(Machiavellianism)	1.67***	--	--	-4.21	-2.18	-0.04	1.97	1.06***	--	--	-2.14	-0.52	1.25	2.82
9. Most people...(Machiavellianism)	1.28***	--	--	-3.07	-1.10	0.23	2.04	0.87***	--	--	-2.58	-0.55	0.78	2.03
10. People see me...(Narcissism)	--	1.17***	--	-4.78	-2.38	-0.22	1.55	--	1.72***	--	-4.17	-2.44	-0.28	1.68
11. I hate being the...(R; Narcissism)	--	0.57***	--	-2.43	-0.89	0.70	2.17	--	0.45***	--	-2.58	-1.09	0.49	1.67
12. Many group activities...(Narcissism)	--	1.48***	--	-5.64	-3.52	-1.20	1.30	--	1.36***	--	-3.38	-1.82	0.19	2.36
13. I know that I am special...(Narcissism)	--	1.60***	--	-6.21	-3.70	-1.14	1.23	--	1.23***	--	-3.34	-1.87	-0.05	1.74
14. I like to get acquainted...(Narcissism)	--	1.29***	--	-3.31	-1.21	0.76	2.49	--	0.84***	--	-2.35	-0.29	1.21	2.57
15. I feel embarrassed...(R; Narcissism)	--	0.48***	--	-2.42	-1.00	-0.04	2.25	--	0.57***	--	-2.70	-1.27	-0.17	1.47

16. I have been compared to...(Narcissism)	--	1.25***	--	-3.95	-1.74	-0.72	1.18	--	1.06***	--	-2.56	-0.71	0.35	1.92
17. I am an average person (R; Narcissism)	--	0.79***	--	-2.88	-1.52	-0.32	1.66	--	0.98***	--	-2.58	-1.62	-0.50	0.89
18. I insist on getting...(Narcissism)	--	0.62***	--	-2.11	0.24	1.70	2.79	--	0.33**	--	-1.03	0.68	2.15	3.53
19. I like to get revenge...(Psychopathy)	--	--	1.45***	-3.50	-1.95	-0.14	1.54	--	--	1.25***	-2.71	-1.22	0.64	2.25
20. I avoid dangerous...(R; Psychopathy)	--	--	0.22*	-2.38	-1.15	-0.02	1.46	--	--	0.59***	-2.32	-0.93	0.40	1.88
21. Payback needs to be...(Psychopathy)	--	--	1.20***	-4.49	-2.85	-1.30	0.42	--	--	1.00***	-2.62	-1.27	0.04	1.36
22. People often say...(Psychopathy)	--	--	1.38***	-4.46	-2.53	-1.27	0.48	--	--	0.77***	-2.57	-0.83	0.31	1.48
23. It's true that I can be...(Psychopathy)	--	--	1.28***	-3.23	-1.04	0.22	1.83	--	--	0.94***	-2.32	-0.40	0.94	2.33
24. People who mess with...(Psychopathy)	--	--	1.74***	-4.01	-2.46	-0.53	1.60	--	--	1.17***	-2.67	-1.08	0.76	2.47
25. I have never gotten...(R; Psychopathy)	--	--	0.63***	-3.14	-2.21	-1.62	-0.42	--	--	0.35**	-2.18	-1.33	-0.81	0.13
26. I enjoy having sex...(Psychopathy)	--	--	1.08***	-3.68	-2.59	-1.54	-0.30	--	--	0.70***	-2.26	-1.47	-0.52	0.34
27. I'll say anything to get...(Psychopathy)	--	--	1.45***	-5.17	-3.55	-1.98	0.16	--	--	1.08***	-2.42	-1.25	0.19	1.53

Note. G: General dark triad factor; Mach: Machiavellianism factor; Narc: Narcissism factor; Psych: Psychopathy factor; *a1-a4*: item discrimination estimates on the general and specific latent dimensions; *d1-d4*: item threshold (i.e., difficulty) estimates; --: item parameter fixed at zero. *: $p < .05$; **: $p < .01$; ***: $p < .001$

Table 5.

Correlations of Residualized SD3 Scores with Criterion Measures in Adult and Adolescent Samples.

Sample 1 (University Students; $N=678$)	Sample 2 (Adolescent High School Students; $N=442$)
--	--

SD3 subscales	Mach IV	FFNI-SF GN	TriPM Total Score	MACH IV	FFNI-SF GN	TriPM Total Score
Machiavellianism	.45*^a	.25* ^b	.13* ^b	.33*^a	.16* ^b	.18* ^b
Narcissism	-.12* ^b	.34*^a	-.01 ^b	.06 ^b	.31*^a	.14* ^b
Psychopathy	.16* ^b	.11* ^b	.31*^a	.13 ^b	.18* ^b	.35*^a
<i>M</i>	2.78	101.41	67.11	2.86	114.96	69.00
<i>SD</i>	0.43	20.68	27.10	0.40	21.95	20.64
<i>α</i>	.80	.90	.93	.63	.87	.89

Note. Mach IV = Machiavellianism Inventory-Version IV; FFNI-SF = Five Factor Narcissism Inventory-Short Form; GN = Grandiose Narcissism scale; TriPM: Triarchic Psychopathy Measure. In each sample, the nominal significance level (i.e., $p < .05$) was corrected according to the Bonferroni procedure and set at $p < .0025$. Bold highlights convergent validity coefficients. Within each row and sample, different superscripts indicate that convergent validity coefficients were significantly different ($p < .05$) from discriminant validity coefficients according to Steiger's z test.

* $p < .0025$

Figure 1.

Rest-score functions for SD3 Items in Sample 1 ($N=678$). Items in order (1–27) from left to right.



Figure 2.

Rest-score functions for SD3 Items in Sample 2 ($N=442$). Items in order (1–27) from left to right.



Electronic Supplementary Material

Evaluating the Psychometric Properties of the Short Dark Triad (SD3)

in Italian Adults and Adolescents

Methods

Participants

Sample 2. This sample was composed of 599 adolescents attending large public high schools in Central Italy. Of these, 301 participants (50.3%) were female, and 295 (49.2%) were male, with 3 (0.5%) refusing to disclose their gender. Their mean age was 16.1 years, $SD = 1.47$ years, although three (0.5%) refused to report their age. One hundred fifty-seven adolescents (26.2%) reported incomplete data on one or more measures. Questionnaires were considered incomplete if more than 10% of the items in any given scale were not answered: These participants were excluded from the final sample. Notwithstanding the non-negligible proportion of participants who provided incomplete data, missing values were completely at random, Little's MCAR $\chi^2(93) = 82.16, p > .70$.

Adolescents who reported incomplete questionnaires did not differ significantly from adolescents who reported complete measures on gender, $\chi^2(1) = 0.38, p > .50, \phi = .03$, age, $t(594) = 1.67, p > .05, d = 0.14$, frequency of school failure, $\chi^2(1) = 0.18, p > .60, \phi = .02$, and conduct grade, Mann-Whitney $U = 31203.00, z = -0.98, p > .30$, rank biserial $r = -.04$. However, adolescents with missing data on one or more questionnaires had significantly (albeit weakly) lower average school grade ($M = 6.52, SD = 1.01$) than adolescents with complete data on all questionnaires ($M = 6.78, SD = 0.90$), Mann-Whitney $U = 28373.50, z = -2.61, p < .01$, rank biserial $r = -.13$.

Translation Procedures

Equivalence with the original meaning of the items was the guiding principle in the translation process (Denissen, Geenen, van Aken, Gosling, & Potter, 2008). First, the English SD3 was independently translated into Italian by one of the authors (A.F.), and by two other clinical psychologists who were fluent in English. After reaching a consensus, we had a native English professional translator translate the Italian version back into English. Finally, this English back-

translation (Cha, Kim & Erlen, 2007; Van De Vijver & Hambleton, 1996; Geisinger, 1994) was sent to an author of the SD3. If the latest version differed from the English original, the translators came to an agreement on the definitive Italian translation. The authors followed the same procedure for translating the TriPM, FFNI-SF, and Mach-IV.

Procedure

To qualify for the present study, participants were required to speak Italian as their first language, a step designed to avoid cultural and lexical bias in questionnaire responses. None of the participants received an incentive, either directly or indirectly, for participating. The questionnaires were administered in random order and anonymously in individual sessions by graduate research assistants. In Sample 1, adult participants were asked to sign a written informed consent form to take part in the study. All participants were at least 18 years old at the time the study was carried out.

In Sample 2, participants gave their written consent to participate in the study after it had been explained to them. When participants were of minor age (i.e., below age 18), parents also had to sign written informed consent forms to allow participation. After obtaining Institutional Review Board approval from the University and the principals of the schools, researchers recruited subjects from university classrooms. The questionnaires were administered in random order and anonymously during class time in school by graduate research assistants when teachers were not present in the classrooms.

Data analyses

Item analyses were carried out according to a multi-trait paradigm (Nunnally & Bernstein, 1994), which involved the computation of two distinct item-total correlation coefficients for each SD3 item. Item *convergent validity* was assessed through computation of item-total correlations corrected for item-total overlap (r_{i-t}) between each item and the total score of the scale to which the item was assigned. Item *discriminant validity* was evaluated by correlating each item with the total score for the scale to which it was not assigned. Multi-trait analysis determines the extent to which

items correlate more strongly with their own domain than with other domains in the instrument.

Differences in excess of twice the standard error of the correlation coefficient — i.e., $2 * (1/\sqrt{N-3})$ — indicate the strength of the association (e.g., Virues-Ortega, Segui-Duran, Desqualzo-Quero, Carnerero, & Martin, 2010). In our data, the critical values for a relevant difference were .08 and .10 in Sample 1 and Sample 2, respectively. Internal consistency reliabilities for all self-report measures used in our study were estimated with Cronbach α coefficients and mean inter-item correlations (MIC).

In the present study, we relied on confirmatory full-information maximum likelihood IRT factor analysis based on graded response model (Samejima, 1997) to estimate the latent structure of SD3 items. Based on Reise and colleagues' (2007) guidelines, we fitted the following three models: a) unidimensional model; b) three-factor model with orthogonal factors, in which items were forced to load on the corresponding factor they were a priori assigned to; and c) three-factor model with correlated factors, in which items were forced to load on the corresponding factor they were a priori assigned to.

Goodness of fit of confirmatory IRT models was evaluated using a) M_2^* statistic; b) root mean square error of approximation (RMSEA) and 90% confidence interval; RMSEA₂ value, I.E., RMSEA value based on M_2^* (Mayedu-Olivares, 2015); c) standardized root mean residual (SRMSR); d) Akaike information criterion (AIC); e) Bayesian information criterion (BIC); and f) sample-adjusted BIC (SABIC) (Cai & Hansen, 2013; Cai, Maydeu-Olivares, Coffman, & Thissen, 2006; Reise, 2012). Although the generalizability of structural equation modeling fit benchmarks to an IRT context may be questionable (e.g., Reise & Rodriguez, 2016), limited Monte Carlo data are available for RMSEA in bifactor IRT models (Cai & Hansen, 2013). Maydeu-Olivares (2015) recently proposed RMSEA₂ $\leq$ 0.089 and SRMR $\leq$.050 for adequate model fit, RMSEA₂ $\leq$ 0.050 and SRMR $\leq$.027 for close fit, and RMSEA₂ $\leq$ 0.050/($k-1$) and SRMR $\leq$.027/($k-1$), where k is the number of response alternatives, for excellent fit in polytomous item IRT factor analyses.

In the present study, we formally tested the latent monotonicity assumption (Reise & Rodriguez, 2016; van der Ark, 2007) with respect to narcissism, psychopathy, and Machiavellianism latent dimensions, respectively, as well as with respect to a latent dimension that was hypothesized to be common to all SD3 items. This assumption was checked through the examination of rest-score functions (Meijer et al. 2015). Rest-score functions are simply plots of the raw-total score (minus the item score) and the item proportion endorsed within a given rest-score grouping. In all statistical analyses the level for statistical significance was $p < .05$; in the case of multiple comparisons the nominal significance level was corrected according to the Bonferroni procedure for multiple comparisons.

Consistent with past studies that have validated the SD3 with concurrent criteria (Jones & Paulhus, 2014), nomological network validity was evaluated assessing the associations between SD3 residualized measures of Machiavellianism, narcissism and psychopathy, and the Mach IV total score, the FFNI-SF GN scale scores and the TriPM total scores, respectively. Significance for the difference between two overlapping correlations (i.e., correlations with one variable in common) was tested through the computation of Steiger's (1980) z -test.

Full-information maximum likelihood IRT factor analysis models and indices were estimated using 'mirt' R package (Chalmers et al., 2012), rest-score functions were computed using 'mokken' R package (van der Ark, 2014). Steiger's (1980) z test for the difference between two overlapping correlation coefficients was computed using 'cocor' R package (Diedenhofen & Musch, 2015). All other analyses were computed using SPSS 22. For completeness, we have included detailed item analyses throughout.

References

- Cai, L., & Hansen, M. (2013). Limited-information goodness-of-fit testing of hierarchical item factor models. *British Journal of Mathematical and Statistical Psychology*, 66, 245-276.
- Cai, L., Maydeu-Olivares, A., Coffman, D. L., & Thissen, D. (2006). Limited information goodness-of-fit testing of item response theory models for sparse 2p tables. *British Journal of Mathematical and Statistical Psychology*, 59, 173-194.
- Cha, E. S., Kim, K. H., & Erlen, J. A. (2007). Translation of scales in cross-cultural research: Issues and techniques. *Journal of Advanced Nursing*, 58, 386–395.
- Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory package for the R environment. *Journal of Statistical Software*, 48, 1-29.
- Denissen, J. J., Geenen, R., van Aken, M. A., Gosling, S. D., & Potter, J. (2008). Development and validation of a Dutch translation of the Big Five Inventory (BFI). *Journal of Personality Assessment*, 90, 152–157.
- Diedenhofen, B. & Musch, J. (2015). cocor: A Comprehensive solution for the statistical comparison of correlations. *PLoS ONE*, 10, e0121945.
- Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, 6, 304–312.
- Maydeu-Olivares, A. (2015). Evaluating the fit of IRT models. In S. P. Reise & D. A. Revicki (Eds.), *Handbook of item response theory modeling: Applications to typical performance assessment* (pp. 111-127). New York, NY: Routledge.
- Meijer, R.R., Tendeiro, J.N., Wanders, R.B.K. (2015). The use of nonparametric item response theory to explore data quality. In S. P. Reise and D. A. Revicki (Eds.), *Handbook of Item response Theory modeling: Applications to typical performance assessment* (pp. 85–110). New York, NY: Routledge.

- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- Reise, S. P., Morizot, J., & Hays, R. D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research, 16*, 19–31.
- Reise, S. P., & Rodriguez, A. (2016). Item response theory and the measurement of psychiatric constructs: some empirical and conceptual issues and challenges. *Psychological Medicine, 46*, 2025-2039.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*, 245-251.
- van de Vijver, F. J. R., & Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist, 1*, 89–99.
- van der Ark, L. A. (2007). Mokken scale analysis in R. *Journal of Statistical Software, 20*, 1-19.
- van der Ark, L.A. (2014). New developments in Mokken scale analysis in R. *Journal of Statistical Software 48*, 1–27.
- Virúes-Ortega, J., Seguí-Duran, D., Descalzo-Quero, A., Carnerero, J. J., & Martín, N. (2011). Caregivers' agreement and validity of indirect functional analysis: A cross-cultural evaluation across multiple problem behavior topographies. *Journal of Autism and Developmental Disorders, 41*, 82-91.