The (mis)measurement of the Dark Triad Dirty Dozen: Exploitation at the core of the scale

Petri Kajonius, Björn Persson, Patricia Rosenberg, Danilo Garcia

Background: The dark side of human character has been conceptualized in the Dark Triad Model: Machiavellianism, psychopathy, and narcissism. These three dark traits are often measured using single long instruments for each one of the traits. Nevertheless, there is a necessity of short and valid personality measures in psychological research. As an independent research group, we replicated the factor structure, convergent validity and item response for one of the most recent and widely used short measures to operationalize these malevolent traits, namely, Jonason's Dark Triad Dirty Dozen. We aimed to expand the understanding of what the Dirty Dozen really captures because the mixed results on construct validity in previous research.

Method: We used the largest sample to date to respond to the Dirty Dozen (N = 3,698). We firstly investigated the Dirty Dozen's factor structure using Confirmatory Factor Analysis. Secondly, using sub-sample (n = 500) and correlation analyses, we investigated the Dirty Dozen dark traits convergent validity to Machiavellianism measured by the Mach-IV, psychopathy measured by Eysenck's Personality Questionnaire Revised, narcissisms using the Narcissism Personality Inventory, and both neuroticism and extraversion from the Eysenck's questionnaire. Finally, besides these Classic Test Theory analyses, we analyzed the responses for each Dirty Dozen item using Item Response Theory (IRT).

Results: The results confirmed previous findings of a bi-factor model fit: one latent core dark trait, plus the three dark traits. An additional exploratory distribution analysis showed that all three Dirty Dozen traits had a striking bi-modal distribution, which might indicate unconcealed social undesirability with the items. The three Dirty Dozen traits did converge to, although not strongly, with the contiguous single Dark Triad scales (r between .41-.49). The probabilities of filling out steps on the Dirty Dozen narcissism-items were much higher than on the Dirty Dozen items for Machiavellianism and psychopathy. Overall, the Dirty Dozen instrument delivered the most predictive value with persons with average and high Dark Triad traits (Theta > -0.5). Moreover, the Dirty Dozen scale was better conceptualized as measured of a combined Machiavellianism-psychopathy factor, not narcissism, that can

be replaced with item 4: 'I tend to exploit others towards my own end'.

Conclusion: The Dirty Dozen showed a consistent factor structure, a relatively convergent validity similar to that found in earlier studies. Narcissism measured using the Dirty Dozen, however, did not contribute with information to the core constitution of the Dirty Dozen construct. More importantly, the results imply a Single Item Dirty Dark Triad (SIDDT) measure of a manipulative and anti-social core as the content of the Dirty Dozen scale.

1

2

3

4

5

6

7

8

9   The (mis)measurement of the Dark Triad Dirty Dozen: Exploitation at the core of the scale

10   Petri J. Kajonius[1, 2, 3], Björn N. Persson[3, 4], Patricia Rosenberg[3], Danilo Garcia[1, 3, 5, 6]*

11   [1]Department of Psychology, University of Gothenburg, Sweden

12   [2]Department of Social Psychology, University of Skövde, Sweden

13   [3]Network for Empowerment and Well-Being, Gothenburg, Sweden

14   [4]Department of Cognitive Neuroscience and Philosophy, University of Skövde, Sweden

15   [5]Blekinge Center of Competence, Blekinge County Council, Karlskrona, Sweden

16   [6]Centre for Ethics, Law, and Mental Health (CELAM), University of Gothenburg, Sweden

17

18

19

20

21

22

23 * Correspondence concerning this article should be addressed to D. Garcia, Network for

24 Empowerment and Well-Being, Axel W. Anderssons väg 8A, SE 371 62 Lyckeby, Sweden. E-

25 mail: danilo.garcia@icloud.com.

26   **Abstract**

27 **Background:** The dark side of human character has been conceptualized in the Dark Triad
28 Model: Machiavellianism, psychopathy, and narcissism. These three dark traits are often
29 measured using single long instruments for each one of the traits. Nevertheless, there is a
30 necessity of short and valid personality measures in psychological research. As an independent
31 research group, we replicated the factor structure, convergent validity and item response for one
32 of the most recent and widely used short measures to operationalize these malevolent traits,
33 namely, Jonason's Dark Triad Dirty Dozen. We aimed to expand the understanding of what the
34 Dirty Dozen really captures because the mixed results on construct validity in previous research.
35
36 **Method:** We used the largest sample to date to respond to the Dirty Dozen (N = 3,698). We
37 firstly investigated the Dirty Dozen's factor structure using Confirmatory Factor Analysis.
38 Secondly, using sub-sample (n = 500) and correlation analyses, we investigated the Dirty Dozen
39 dark traits convergent validity to Machiavellianism measured by the Mach-IV, psychopathy
40 measured by Eysenck's Personality Questionnaire Revised, narcissisms using the Narcissism
41 Personality Inventory, and both neuroticism and extraversion from the Eysenck's questionnaire.
42 Finally, besides these Classic Test Theory analyses, we analyzed the responses for each Dirty
43 Dozen item using Item Response Theory (IRT).
44
45 **Results:** The results confirmed previous findings of a bi-factor model fit: one latent core dark
46 trait, plus the three dark traits. An additional exploratory distribution analysis showed that all
47 three Dirty Dozen traits had a striking bi-modal distribution, which might indicate unconcealed
48 social undesirability with the items. The three Dirty Dozen traits did converge to, although not
49 strongly, with the contiguous single Dark Triad scales (r between .41-.49). The probabilities of
50 filling out steps on the Dirty Dozen narcissism-items were much higher than on the Dirty Dozen
51 items for Machiavellianism and psychopathy. Overall, the Dirty Dozen instrument delivered the
52 most predictive value with persons with average and high Dark Triad traits (Theta > -0.5).
53 Moreover, the Dirty Dozen scale was better conceptualized as measured of a combined
54 Machiavellianism-psychopathy factor, not narcissism, that can be replaced with item 4: 'I tend to
55 exploit others towards my own end'.
56
57 **Conclusion:** The Dirty Dozen showed a consistent factor structure, a relatively convergent
58 validity similar to that found in earlier studies. Narcissism measured using the Dirty Dozen,
59 however, did not contribute with information to the core constitution of the Dirty Dozen
60 construct. More importantly, the results imply a Single Item Dirty Dark Triad (SIDDT) measure
61 of a manipulative and anti-social core as the content of the Dirty Dozen scale.
62
63 *Keywords*: Dark Triad; Dark Triad Dirty Dozen; Gender; Item Response Theory;
64 Machiavellianism; Narcissism; Psychopathy; Single Item Dirty Dark Triad.

65
66
67
68
69
70
71
72
73
74   Over the last 25 years, the vast majority of personality research has focused on the Big

75   Five traits: openness, conscientiousness, extraversion, agreeableness, and neuroticism. The Big

76   Five Model of personality is a theory developed from both language taxonomy as well as

77   statistical factor analysis (Costa & McCrae, 1992). However, the Big Five has painted an

78   unnecessarily "light" conceptualization of human nature (Lee & Ashton, 2014). Critics have also

79   argued against − what they believe is − the overreliance on factor analysis (i.e., one of the

80   methods in Classical Test Theory) to uncover the latent structure of personality, without a well-

81   grounded theoretical basis and substantial variation in methodology (e.g., Block, 1995; Gould,

82   1981). In the last decade, personality psychologists have turned their attention to the dark side of

83   human character: Machiavellianism, psychopathy, and narcissism. Together, these traits are

84   widely known as the Dark Triad model (Paulhus & Williams, 2002). The validation studies of

85   Dark Triad measures have mostly been conducted using Classic Test Theory methods and in

86   very few cases using Item Response Theory (IRT) methods. The Dark Triad embodies

87   interpersonal, sub-clinical, and maladaptive personality traits in the general population (Paulhus

88   & Williams, 2002), which are characterized by manipulativeness (i.e., Machiavellianism),

89   impulsivity and antagonism (i.e., psychopathy), and the sense of entitlement (i.e., narcissism).

90   The Dark Triad traits are associated with a value system of unconventional and antisocial

91   morality (Furnham, Richards, & Paulhus, 2013; Kajonius, Persson, & Jonason, 2015). In

92   essence, individuals with high levels on any of these dark traits appear to operate in selfish and

93   competitive ways with a common core: uncooperativeness (see Jones & Figueredo, 2013). Thus,

94   whether the dark traits constitute a ternary model of unique traits or a unified uncooperative

95   general factor with three closely related anti-social sub-traits is still an open question. In this

96   context, validation studies using IRT methods might shade some light in what different measures

97   of the Dark Triad actually measure.

98   As with most personality psychology research, the measurement of individuals' tendencies

99   on the dark traits is often conducted using self-report measures. Most of the time, this has been

100   done using one instrument for each trait. These single instruments to measure the dark traits are

101   often long and time demanding. For the trait of Machiavellianism, for example, researchers often

102   use Christie and Geis' Mach-IV (1970), which was originally based on statements from the

103   Italian Niccolò Machiavelli´s books *The Prince* and *The Discourse* (see also Jones & Paulhus,

104   2009, who point out that the instrument also captures behaviors from the Chinese military

105   general, strategist, and philosopher Sun Tzu's book *The Art of War*; behaviors such as planning,

106   building a reputation, and creating alliances). The trait of psychopathy is often measured with the

107   Self-report Psychopathy Scale (Hare, 1985). This instrument was first used on prisoners and later

108   on validated in non-criminal populations as well (Hare, 1985). Nevertheless, also the

109   psychoticism scale in the hierarchical three-factor model proposed by Eysenck (e.g., Eysenck,

110   Eysenck, & Barrett 1985) has been used as a measure of psychopathy or "Impulsive

111   Unsocialized Sensation Seeking" (Zuckerman, Kuhlman, Thornquist, & Kiers, 1991. See also

112   Zuckerman, 1989, 1991; Linton & Power, 2013; Garcia & Sikström, 2014). Finally, narcissism is

113   often measured using the Narcissism Personality Inventory, which comprises 80 (long version)

114   or 32 (short version) paired-items (Raskin & Hall, 1979). Nevertheless, shorter measures

115   comprising all three traits in one single instrument have been created to facilitate data collection.

116   One such measure is the Dark Triad Dirty Dozen (Jonason & Webster, 2010). See Table

117   1 for the statements and the key word in each one of the statements in the Dirty Dozen scale. The

118   Dirty Dozen comprises 12 items that in four studies were demonstrated to retain its core of

119   disagreeableness when compared to 91 items from questionnaires that measured the dark traits

120   separately (Jonason & Webster, 2010). This is a reduced item count by 87%. Subsequent studies

121   with smaller samples have explored the thin line between efficiency and accuracy in this short

122   scale. The findings suggest a bi-factor structural model with both a general latent Dark Triad

123   construct and the three dark traits of Machiavellianism, psychopathy, and narcissism. In addition,

124   the findings also show relatively good convergent validity with the Mach-IV ($r = .53$), Self-

125   report Psychopathy Scale III ($r = .32$), and Narcisssim Personality Inventory-40 ($r = .53$)

126   (Jonason & Luévano, 2013). Further validations, using a sample of young undergraduates, were

127   reported with the ubiquitous Big Five Inventory developed by Benet-Martínez and John (1998).

128   The findings revealed an unstable core of conscientiousness for psychopathy and agreeableness

129   for both Machiavellianism and psychopathy, with no clear, correlational relationships with

130   extraversion for narcissism (Jonason, Kaufman ,Webster, & Geher, 2013). Again, a bi-factor

131   model (i.e., one general factor plus three specific factors) fit the data best. This suggests that each

132   dark trait measured something unique (Jonason, Kaufman, Webster, & Geher, 2013), in addition

133   to the common variance captured by the general factor. In addition to validations using Classical

134   Test Theory, current research has moved to IRT for the validation of the Dirty Dozen scale.

135   <u>Table 1 should be here</u>

136   There is a large diversity of models that have been developed using IRT. For simplicity

137   reasons, we refer to most of them using the global term IRT throughout the rest of the paper. IRT

138   was first proposed in the field of psychometrics for the purpose of ability assessment. For

139   instance, all major educational tests are developed using this technique because it significantly

140   improves measurement accuracy and reliability, and it provides significant reductions in

141   assessment time and effort (for a review see An & Yung, 2014). In recent years, this technique

142   has also been applied in health and clinical research (e.g., Hays, Morales, & Reise, 2000; Edelen

143   & Reeve, 2007; Holman, Glas, & de Haan, 2003; Reise & Waller, 2009). Using IRT models

144   researchers have found a slightly lower endorsement threshold of the dark traits for males

145   compared to females. This has been interpreted as differences in social undesirability sensitivity,

146   or true differences as proposed by mating-strategy theory (Webster & Jonason, 2013). The latest

147   validation study among onsite UK undergraduates and online Crowdflower-workers[1], however,

148   found conflicting results using Mokken analysis, a non-parametric form of IRT (Carter,

149   Campbell, Muncer, & Carter, 2015). While the expected three traits of Machiavellianism,

150   psychopathy, and narcissism, emerged among female students' scores; only two traits emerged

151   among male students' scores. These two traits were a combined Machiavellianism-psychopathy

152   factor and a narcissism factor. In contrast, among the online workers, only one core construct of

153   the Dirty Dozen appeared. These differences were not explained by invariance over sex and age.

154   Hence, casting some uncertainty on the evasive constructs measured by the Dirty Dozen scale or

155   suggesting some kind of mismeasurement of the triad by this specific scale.

156   **The Present Study**

157   The possibility to replicate findings is one of the parameters that distinguish science from non-

158   science. In short, replication should be at "the heart of science" (Schmidt, 2009). By use of

159   conceptual replications we can potentially confirm which findings about human nature that can

---

[1] "CrowdFlower is a data enrichment, data mining and crowdsourcing company based in the Mission District of San Francisco, California. The company's software as a service platform allows users to access an online workforce of millions of people to clean, label and enrich data. CrowdFlower is typically used by data scientists at academic institutions, start-ups and large enterprises."
Retrieved from: https://en.wikipedia.org/wiki/CrowdFlower.

160   be generalized and thus increase predictive validity in our regular use of psychological

161   measurements. As researchers we expect that replication studies are common and that the

162   methodology is well developed, however, particularly in social sciences the contrary is true,

163   demonstrating an overall replication rate of only 1.07% (Makel, Plucker, & Hegarty, 2012; see

164   also Lucas & Donnellan, 2013, and the Registered Replication Reports initiative by the

165   Association   for   Psychological   Science,

166   http://www.psychologicalscience.org/index.php/replication).

167   A major problem in current validations is the small sample sizes and a general lack of

168   power and precision. This "results in lower precision in parameter estimates and systematically

169   inflated effect size estimates" (Lucas & Donnellan, 2013, p. 453). In addition, the current

170   validation studies that have been published provide many statistically significant low-powered

171   findings even within the same study, which "paradoxically provide less support for a

172   phenomenon than papers that report some failures to reach statistical significance" (Lucas &

173   Donnellan, 2013, p. 453; see also Francis, 2012; Schimmack, 2012). To the best of our

174   knowledge, the present study provides the largest single sample used to this date ($N = 3,698$) to

175   replicate some of the most common findings with regard to the Dirty Dozen scale. For instance,

176   previous validation studies on the Dirty Dozen have had limited, or at least unclear,

177   generalizability, often only including undergraduates or homogenous age cohorts. In addition,

178   although we do believe in researchers' capacity for objectivity, we see as an important venue that

179   an independent research group that had no ties to the construction of the Dirty Dozen scale

180   conducted the present replication study.

181   In sum, we present a replication of Confirmative Factor Analysis (CFA), Convergence

182   Analyses, and IRT Analyses of the Dirty Dozen, which is one of the most recent popular

183   personality short scales used among personality psychologists to measure the dark traits (see a

184   review by Furnham, Richards, Rangel, & Jones, 2014). We have also conducted new analyses.

185   First, we investigated the original factor structure, which has shown varied results in previous

186   studies, this time using a sizable, heterogeneous sample from all walks and ages of life. For

187   instance, with the large sample at hand, there were sufficient respondents for an exploratory

188   distribution analysis, which has not been reported before. Second, we further establish the

189   validity of the traits measured using the Dirty Dozen by investigating convergence with known,

190   contiguous single long scales of the dark traits: the Mach-IV, Eysenck's Personality

191   Questionnaire Revised, and the Narcissism Personality Inventory. Third, using IRT, we explored

192   what the Dirty Dozen truly endeavors to measure. We contend that what makes the Dark Triad,

193   measured by this specific scale, "dark" is not a uniform stable core but, instead, a challenging

194   mix of malevolent and tradition-laden anti-social and uncooperative traits. If so, we might be

195   able to clarify what the Dirty Dozen measures and uncover what scale-items might be

196   responsible for the many interpretations of its core.

197   **Method**

198   **Ethical statement**

199   After consulting with the Network for Empowerment and Well-Being's Review Board we

200   arrived at the conclusion that the design of the present study (e.g., all participants' data were

201   anonymous and will not be used for commercial or other non-scientific purposes) required only

202   informed consent from the participants.

203   **Participants**

204   The participant data was collected through Mechanical Turk (MTurk), which has demonstrated

205   reliability and validity, providing a wider range of socio-economic backgrounds compared to

206   other samples (Casler, Bickel, & Hackett, 2013). This is particularly useful when it comes to

207   research on values, such as the undesirability of the Dark Triad traits' values (cf. Kajonius et al.,

208   2015). All participants were informed that the survey was voluntary, anonymous, and that the

209   participants could terminate the survey at any time. The MTurk workers received 50 cents (US-

210   dollars) as compensation for participating and only residents of the US were allowed to accept

211   participation. Two control questions were added to the survey, to control for automatic responses

212   (e.g., "This is a control question, please answer "neither agree or disagree"). A total of 50

213   participants responded erroneously to one or both of the control questions, the final sample

214   constituted 3,698 (M_{age} = 33.5, SD = 11.8). As expected, males (N_{males} = 1,726) scored higher on

215   all Dark Triad traits than females ($N_{females}$ = 1,972), as summarized in the descriptive Table 2. A

216   subsample (N = 500) also answered to single long instruments of the Dark Triad, extraversion,

217   and neuroticism.

218   **Measures**

219   The Dark Triad Dirty Dozen (Jonason & Webster, 2010) is a 12-item self-report

220   questionnaire measurement of the three Dark Triad traits. Participants are asked to rate how

221   much they agreed (1 = *Strongly disagree;* 7 = *Strongly agree*) with statements such as: "I tend to

222   manipulate others to get my way" (Machiavellianism), "I tend to lack remorse" (psychopathy),

223   and "I tend to want others to admire me" (narcissism). Items were averaged to create each

224   dimension (Cronbach's Alphas between .74 to .85; see Table 2 for Alphas for both males and

225   females). We also constructed a composite score of the three dark traits by using the mean values

226   from all of the items.[2] For facilitating readability, all measures of the dark traits using the Dirty

[2] We are well aware of the controversy concerning composite scores present in the literature (e.g., Glenn & Sellbom, 2015). While we're inclined to agree with their arguments on a theoretical level, we've elected to use a composite score as it is a quick abbreviation of a general "dark personality". Furthermore, we conducted an exploratory omega analysis (omega in R package *psych*, see also Revelle & Wilt, 2013), which yielded a $\omega_{hierarchical}$ coefficient of .72, which suggests that the Dirty Dozen is saturated by a general factor. Additionally, the correlation between the

227   Dozen are labeled as follows: DD Machiavellianism, DD psychopathy, and DD narcissism. High

228   scores represent high degree in each of the dark traits or, in the case of the composite, a high

229   degree of the Dark Triad Dirty Dozen core.

230   <u>Table 2 should be here</u>

231   The Mach-IV (Christie & Geis, 1970) was used to also measure Machiavellianism. The

232   Mach-IV consists of 20 items that reflect ways of thinking and opinions about people and

233   different situations (e.g., "Never tell anyone the real reason you did something unless it is useful

234   to do so"). Participants were requested to rate to what extent they agree with each statement on a

235   6-point Likert scale: 1 = *Strongly agree*, 6 = *Strongly disagree*. The Machiavellianism score was

236   computed by summarizing the means across the 20 items, a high score representing high degree

237   of Machiavellianism.

238   The short version of the Eysenck's Personality Questionnaire Revised was used to

239   measure extraversion (e.g., "Do you usually take the initiative in making new friends?"),

240   neuroticism (e.g., "Do you ever feel 'just miserable' for no reason?"), and psychoticism (e.g.,

241   "Would you like other people to be afraid of you?") (Eysenck, Eysenck, & Barrett, 1985). The

242   Eysenck questionnaire consists of 12 items for each trait (forced binary answers: *Yes* or *No*). The

243   score for each of the personality traits was computed as the sum of the 12 items, with *yes*

244   responses coded as 1 and *no* responses coded as 0. Thus, a high score represents high degree in

245   each of the three personality traits. As stated in the Introduction section, Eysenck's psychoticism

246   scale is better labeled as psychopathy (Zuckerman, 1989, 1991). Hence, for the rest of the paper

247   we refer to the psychoticism scale as psychopathy.

---

composite score and an unrotated principal component was .994. These analyses are available from the corresponding author upon request.

248   The short version of the Narcissistic Personality Inventory was used to also measure

249   Narcissism (Ames, Rose & Anderson, 2006). The instrument consists of 16 pairs of items (one

250   consistent and one inconsistent with narcissistic behavior in each pair) for what participants are

251   instructed to choose, for each pair, one item that comes closest to describing their own feelings

252   and beliefs about themselves. The narcissism score was computed as the sum of the 16 items,

253   with narcissism-consistent responses (e.g., "I really like to be the center of attention") coded as 1

254   and narcissism-inconsistent responses coded as 0 (e.g., "It makes me uncomfortable to be the

255   center of attention"). Thus, a high score represents high degree of narcissism.

256   **Statistical Analysis**

257   As in earlier studies, there was a relatively large skewness in the psychopathy scores and kurtosis

258   in the narcissism scores. This has, however, been shown to not have a negative effect on

259   subsequent statistical analysis when the sample size reaches the thousands (Lumley, Diehr,

260   Emerson, & Chen, 2002). First, using Classic Test Theory, we used CFA for testing two

261   contending models, one with only the latent dark triad core branching into three dark traits, and

262   second, a bi-factor model with the latent dark triad core connecting directly with all items, while

263   the three dark traits connecting only to their respective items. We used Structural Equation

264   Modeling (SEM) in the software Amos v.22 for these calculations. Second, using SPSS v. 22, we

265   conducted convergent correlational analyses with the collected contiguous single dark traits

266   scales and Extraversion and Neuroticism (i.e., Mach-IV, Eysenck's Personality Questionnaire

267   Revised, and the Narcissism Personality Inventory). Third, with the purpose of exploring the

268   Dirty Dozen content, we utilized the much in-demand method of IRT using the R package *MIRT*

269   version 1.10 (Chalmers, 2012) in R version 3.2.1 (R Core Team, 2015). This is a methodology

270   for modeling how test items contribute to one latent, scalable trait. We used a graded response

271   model (analogous to the 2PL for dichotomous items) which basically generates two defining

272   characteristics for each item: a slope coefficient, or discrimination parameter alpha (a), and a

273   discrimination coefficient, or threshold parameter beta (b). The a parameter shows how strongly

274   an item relates to a given latent construct theta (θ; which in this study the Dark Triad core as

275   measured by the Dirty Dozen scale). The a parameter can be analogized as a factor loading,

276   whereas the threshold parameters b_{1-6} relates to the level of the latent trait at which the next

277   highest response category has at least 50% probability of being endorsed. For more information

278   about IRT see Morizot, Ainsworth, and Reise (2007).

279   **Results**

280   The first purpose was to replicate the original factor-structure of the Dirty Dozen using Classic

281   Test Theory. Two CFA-models were tested. The first model, a hierarchical structure, with the

282   Dark Triad core "above" the three dark traits, DD Machiavellianism ($\lambda = .75$), DD psychopathy

283   ($\lambda = .86$), and DD narcissism ($\lambda = .56$), was not optimal ($\chi^2(40) = 1530.24$, $p < .001$) and with non-

284   satisfactory fit indices as well (NFI = .92, CFI = .92, and RMSEA = .10). The second model

285   tested was a bi-factor structure, which proved more successful ($\chi^2(28) = 360.19$, $p < .01$) with

286   sufficient fit indices (NFI = .98, CFI = .98, and RMSEA = .05). The RMSEA of this specific

287   model was slightly better than in previous studies (RMSEA = .07 in Jonason & Luévano, 2013;

288   RMSEA = .06 in Jonason et al., 2013). Furthermore, our model showed that 3 out of the 4 items

289   in the DD narcissism cluster had very weak relationships with the Dark Triad core and that the

290   DD Machiavellianism-items demonstrated the strongest relationships. The full model with all

291   items' regression coefficients is reported in Figure 1.

292   Figure 1 should be here

293   In addition, with the large sample at hand, there were sufficient respondents for an

294   exploratory distribution analysis, which as far as we know has not been reported before. Figure 2

295   depicts a strong bimodality of the distribution (peaks on both Likert-categories 1 and 5) found in

296   all three dark traits. DD Psychopathy showed the largest overrepresentation in the lowest scale-

297   category (Likert-category 1), followed by DD Machiavellianism and last DD narcissism.

298   Females were overrepresented in the lowest scale-category (Likert-category 1) compared to

299   males, $N_{females}$ = 280, N_{males} = 151 (DD Machiavellianism), $N_{females}$ = 486, N_{males} = 280 (DD

300   psychopathy), and $N_{females}$ = 188, N_{males} = 114 (DD narcissism). In small sample-studies,

301   distributions such as these, might strongly affect statistical validity, as well as external validity,

302   indicating strong social undesirability with the items.

303   <u>Figure 2 should be here</u>

304   The second purpose was to analyze convergent validity of the dark traits as measured with

305   the Dirty Dozen scale. We were simply looking for the expected, conjoining relationships

306   between the Dark Triad traits and the contiguous, single long scales of the dark traits and both

307   neuroticism and extraversion. Table 3 summarizes the correlations found, which overall did

308   show relatively weak (all rs < .50) converging relationships (Machiavellianism r = .49;

309   psychopathy r = .41, narcissism r = .47). That is, the three dark traits measured with Dirty Dozen

310   showed that DD Machiavellianism and DD psychopathy showed similar correlations, while DD

311   narcissism related less well with the corresponding scales measured using the single long scales.

312   Additionally, the Dirty Dozen Dark Triad composite (Table 3, row 4) showed the smaller

313   relationships with the dark traits measured using the single long scales (correlations between the

314   Dirty Dozen dark traits and the Dark Triad core were between .75-.88, correlations between the

315   dark traits measured with the single instruments and the Dark Triad core were between .31-53).

316   In addition, there are some discrepancies between the correlations between neuroticism and

317   extraversion and the dark traits depending on how the Dark Triad was measured. For example,

318   while there was a weak significantly positive correlation between DD narcissism and neuroticism

319   ($r = .20$, p < .01), there was no significant correlation to neuroticism when narcissism was

320   measured using the Narcissism Personality Inventory ($r = -.07$). In contrast the relationship

321   between narcissism and extraversion was almost twice as large when narcissism was measured

322   using the single instrument ($r = .40$, $p < .01$) than when measured using the Dirty Dozen ($r = .25$,

323    $p < .01$).

324   Table 3 should be here

325   The third and last purpose was to extend the discussion on the construct of the Dark Triad

326   as measured by the Dirty Dozen scale, using IRT. The results from both the CFA model 1 (NFI =

327   .92) and 2 (NFI = .98) indicated adequate unidimensionality, which is the basic assumption for

328   IRT. We ran a polytomous graded-response model on the 12 Dirty Dozen items, allowing items

329   to load on a latent Dark Triad core. The Total Information Curve reported in Figure 3 shows that

330   the core of the Dark Triad (θ) was revealed in a maximized way only when a participant has

331   close to average levels (-0.5) of this latent trait (see Figure 4 for the Information Curve for each

332   one of the dark traits measured using the Dirty Dozen). Hence, the Dirty Dozen scale functions

333   well for capturing average and higher levels of the core Dark Triad, but not the lower levels. This

334   once again leads to the question what constitutes the dark core, or more specifically to what the

335   Dirty Dozen scale actually measures. See Figure 5 in the supplementary material for Scale

336   Information Curves for each of the 12 items of the Dirty Dozen

337   Figure 3 should be here

338   Figure 4 should be here

339   <u>Figure 5 should be here</u>

340   In Table 4 the items' ability to differentiate (a parameter) between people with similar

341   levels of the same latent trait are ranked, starting with the item yielding the most information

342   (item 4: Exploit). The a-parameter typically ranges from 0.5 to 2.0 in personality scales

343   (Morizot, Ainsworth, & Reise, 2007). As can be seen, three of the four DD narcissism items

344   (item 10: Attend, item 9: Admire, and item 11: Status) contributed least to differentiation

345   between individuals. The difficulties (b) for each item are listed in rows, and reflect the threshold

346   levels of the latent trait necessary to have at least 50% chance of endorsing the next scale-step

347   (e.g., b_1 denotes answering Option 1 vs. 2, 3, 4, 5, 6, 7). The b-parameters are scaled on the same

348   metric as the latent trait (θ) and falls in the range of -3 to +3 SD, thus 0 is approximated to be of

349   average difficulty (at the mid-point of the distribution). At the highest scale-step (b_6), all sub-

350   factor items showed extreme difficulty, close to 3 SD. At the lower end (b_1), the items on DD

351   Machiavellianism (items 1-4) and DD psychopathy (items 5-8), still showed much difficulty,

352   close to average 0, while narcissism (items 9-12), showed much less difficulty, close to -2SD.

353   These results imply that DD narcissism is not contributing as much information to the core

354   constitution of the dark triad construct, since the probability of filling out steps on the DD

355   narcissism-items are much higher than on the DD Machiavellianism and DD psychopathy-items

356   (cf. first, the skewness in distributions in Figure 2 and second, that DD narcissism correlated the

357   least with the contiguous scales in Table 3). In other words, when a respondent does fill out high

358   numbers on scale-items on DD Machiavellianism and DD psychopathy, this rapidly predicts the

359   latent level of participant's core dark personality (e.g., item 1: Manipulate, $a = 2.73$ or item 6:

360   Amoral, $a = 1.91$), but not for DD narcissism (e.g., item 10: Attend, $a = .91$). The separate item

361   information curves are found in Figure 4.

362   The item with the highest a (item 4: Exploit, $a = 3.33$), was of particular interest, due to its

363   superior discriminatory ability compared to the others. We surmised that this item in itself would

364   be able to capture the entire Dark Triad core, as measured by the Dirty Dozen scale. This

365   exploitation-item correlated with the summed Dark Triad ($r = .77$), to the same degree that the

366   three dark traits did, DD Machiavellianism ($r = .88$), DD psychopathy ($r = .75$), and DD

367   narcissism ($r = .76$). When exchanging the summed Dark Triad for the single exploit-item, the

368   internal reliability between the constructs was only marginally lowered: inter-item .61 to .51;

369   Cronbach's Alpha .85 to .80. Finally, comparing the single item with the convergence

370   coefficients of the summed Dark Triad composite (Table 3), the single item performed as well or

371   better as a substitute.

372   Table 4 should be here

373   **Discussion**

374   This was a replication study of the popular and much used Dirty Dozen, based on the largest and

375   most diverse sample to date. All previous research results on the bi-factor structure and

376   convergent validity were confirmed. Concerning the previous varying findings on one-factor, bi-

377   factor, or three-factor solutions (cf. Carter et al., 2015), the large sample in the present study

378   warrants to overall lean towards a bi-factor solution. However, a new (old) problem was brought

379   to the surface with the reporting of a strong bi-modal distribution of all three sub-factors,

380   Machiavellianism, psychopathy, and narcissism. This has not been much emphasized in previous

381   publications and a contribution of the present study is to highlight the scope of this problem and

382   admonish for large samples sizes when researching the Dark Triad, which is known to

383   compensate for unwanted distribution skewness in statistical analyses (cf. Lumley et al., 2002).

384   In the wake of the Mokken analysis by Carter et al. (2015), one of the attempts of this

385   paper was to further the discussion on what the Dark Triad trait consists of and what the Dirty

386   Dozen seeks to measure. First, the distribution analyses (Figure 2) can be interpreted as an inertia

387   to filling out Machiavellianism- and psychopathy-items, while narcissism-items showed normal,

388   unskewed distribution, and consequently, not adding as much to the prediction of the core

389   construct. A second clue to the latent core of the Dark Triad is the convergence analysis in Table

390   3 indicating that narcissism was the sub-factor that least correlated with the adjacent constructs

391   of sub-factors to Mach-IV and EPQ, again showing that Machiavellianism-psychopathy is at the

392   center of the construct. Nevertheless, although we used convergent analyses as in many other

393   validation studies; recent research suggests that short scales should not be validated using these

394   type of analysis (Olaru, Witthöft & Wilhelm, 2015). Third, the IRT-analysis showed that

395   narcissism-items (e.g., need for admiration, attention, and status) had the least difficulty and the

396   least discriminating power, not contributing to the total information on the latent dark trait. We

397   conclude and submit for future research that the Dirty Dozen is a measurement consisting of a

398   core found in Machiavellianism-psychopathy (e.g., manipulation, deceit, amorality, and

399   callousness, with no remorse, as seen in Table 4).

400   Our proposal is that the Dark Triad, at least as measured by the Dirty Dozen, might be the

401   product of a hasty grouping of two "difficult" sub-factors, Machiavellianism and psychopathy,

402   together with one 'easy' sub-factor, narcissism. Being narcissistic is considered more normal in

403   these days and times (Twenge, Campbell, & Freeman, 2012), not as undesirable to fill out in

404   questionnaires, and does not add to the core of the construct. This grouping of three is

405   unfortunate both from a social desirability- (method artifact) and a subclinical perspective (how

406   to find people with real problems). If one wants to quickly find the core of the Dark Triad, a one-

407   item of "I exploit others" might be the proper one-item ultra-brief scale to use. A similar

408   approach has been taken recently with narcissism, compressing the original 40-item scale into a

409   Single Item Narcissism Scale (SINS), and demonstrating sufficient reliability and validity in

410   initial studies (Konrath, Meier, & Bushman, 2014; van der Linden & Rosenthal, 2015).

411   **Future research and conclusions**

412   The varying social undesirability with all sub-factors measured by the Dirty Dozen should

413   be further explored. It is not clear if the bi-modality of distribution is a reflection of this, or if it is

414   a certain group of people with for instance very high Big Five-agreeableness, thus virtually

415   hitting zero on all sub-factors of the Dirty Dozen. In other words, it is not clear if this indicates a

416   genuine difference in the Dark Triad as a construct that is not instrument-specific. Additional

417   studies have to be carried out using different methods and measures in order to assess whether or

418   not such a difference is a method artifact or a real difference.

419   Another problem is that it is not apparent to what extent a short Dark Triad scale taps into

420   and is confounded by clinical populations. In a large replication study such as the present,

421   statistically 1-5% will be eligible for personality disorders. The results from IRT implicates that

422   item-difficulties are sufficient on Machiavellianism and psychopathy to be able to distinguish

423   problematic levels of the dark personality core, but not narcissism.

424   The conclusion on our part is that the Dirty Dozen has its advantages by being short,

425   intuitive, and even fun, containing high face validity, but also has drawbacks by being highly

426   differing in item-difficulties. The mismeasurement of the Dirty Dozen seems to be that it actually

427   measures two constructs, narcissism and an anti-social trait. This specific conclusion is supported

428   by the fact of what we choose to call a Single Item Dirty Dark Triad (SIDDT), the "exploit"

429   item. In situations of restrained research time and space, the SIDDT captures the essence of what

430   the Dirty Dozen actually measures.

431   *"Show me again, the power of the darkness, and I'll let nothing stand in our way. Show me,*

432   *grandfather, and I will finish what you started."*

433   *From Star Wars: The Force Awakens*

434   **References**

435   Ames, D. R., Rose, P., & Anderson, C. P. (2006). The NPI-16 as a short measure of narcissism.

436   *Journal of Research in Personality*, *40*(4), 440-450.

437   An, X., & Yung, Y-F. (2014). Item response theory: What it is and how you can use the IRT

438   procedure to apply it. Paper SAS364-2014.

439   Benet-Martínez, V., & John, O. P. (1998). Los cinco grandes across cultures and ethnic groups:

440   Multitrait multimethod analysis of the Big Five in Spanish and English. *Journal of*

441   *Personality and Social Psychology*, *75*, 729-750.

442   Block, J. (1995). A contrarian view of the five-factor approach to personality description.

443   Psychological bulletin, 117(2), 187–215. doi:10.1037/0033-2909.117.2.187

444   Carter, G. L., Campbell, A. C., Muncer, S., & Carter, K. A. (2015). A Mokken analysis of the

445   Dark Triad 'Dirty Dozen': Sex and age differences in scale structures, and issues with

446   individual items. *Personality and Individual Differences*, *83*, 185-191.

447   Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants

448   and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing.

449   *Computers in Human Behavior*, *29*(6), 2156-2160.

450   Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R

451   environment. *Journal of Statistical Software*, *48*(6), 1-29.

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

453   Costa, P. T., & McCrae, R. R. (1992). Revised NEO personality inventory (NEO PI-R). Odessa,

454   FL: Psychological Assessment Resources.

455   Edelen, M. O., & Reeve, B. B. (2007). Applying item response theory (IRT) modeling to

456   questionnaire development, evaluation, and refinement. *Quality of Life Research*, *16*(1), 5-

457   18.

458   Eysenck, S. B., Eysenck, H. J., & Barrett, P. (1985). A revised version of the psychoticism scale.

459   *Personality and Individual Differences*, *6*(1), 21-29.

460   Francis, G. (2012). Too good to be true: Publication bias in two prominent studies from

461   experimental psychology. *Psychonomic Bulletin & Review, 19*(2), 151–156.

462   Furnham, A., Richards, S.C., & Paulhus, D. L. (2013). The Dark Triad of personality: A 10 year

463   review. *Social and Personality Psychology Compass*, *7*, 199–216.

464   Furnham, A., Richards, S., Rangel, L., & Jones, D. N. (2014). Measuring malevolence:

465   Quantitative issues surrounding the Dark Triad of personality. *Personality and Individual*

466   *Differences*, *67*, 114-121.

467   Garcia, D., & Sikström, S. (2014). The dark side of Facebook: Semantic representations of status

468   updates predict the Dark Triad of personality. *Personality and Individual Differences*, *67*,

469   92-96.

470   Glenn, A. L., & Sellbom, M. (2015). Theoretical and empirical concerns regarding the dark triad

471   as a construct. *Journal of Personality Disorders*, *29*(3), 360-377.

472   Gould, S. J. (1981). *The Mismeasure of Man*. New York: W. W. Norton & Company.

473   Hare, R. D. (1985). Comparison of procedures for the assessment of psychopathy. *Journal of*

474   *Consulting and Clinical Psychology*, *53*(1), 7.

475   Hays, R. D., Morales, L. S., & Reise, S. P. (2000). Item response theory and health outcomes

476   measurement in the 21st century. *Medical Care*, *38*(9), II-28–II-42.

477   Holman, R., Glas, C. A., & de Haan, R. J. (2003). Power analysis in randomized clinical trials

478   based on item response theory. *Controlled Clinical Trials*, *24*(4), 390-410.

479   Jonason, P. K., Kaufman, S. B., Webster, G. D., & Geher, G. (2013). What Lies Beneath the

480   Dark Triad Dirty Dozen: Varied Relations with the Big Five. *Individual Differences*

481   *Research 11*(2), 81–90.

482   Jonason, P. K., & Luévano, V. X. (2013). Walking the thin line between efficiency and accuracy:

483   Validity and structural properties of the Dirty Dozen. *Personality and Individual*

484   *Differences*, *55*(1), 76-81.

485   Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: a concise measure of the dark triad.

486   *Psychological Assessment*, *22*(2), 420-432.

487   Jones, D. N., & Figueredo, A. J. (2013). The core of darkness: Uncovering the heart of the Dark

488   Triad. *European Journal of Personality, 27*, 521-531.

489   Jones, D. N., & Paulhus, D. L. (2009). Machiavellianism. In M. R. Leary & R. H. Hoyle (Eds.),

490   *Handbook of individual differences in social behavior* (pp. 93-108). New York, NY:

491   Guilford Press.

492   Kajonius, P. J., Persson, B. N., & Jonason, P. K. (2015). Hedonism, Achievement, and Power:

493   Universal values that characterize the Dark Triad. *Personality and Individual Differences,*

494   *77*, 173-178.

495   Konrath S., Meier B. P., & Bushman, B. J. (2014). Development and Validation of the Single

496   Item Narcissism Scale (SINS). *PLoS ONE, 9*(8): e0103469.

497   Lee, K., & Ashton, M.C. (2014). The Dark Triad, the Big Five, and the HEXACO model.

498   *Personality and Individual Differences, 67*, 2-5.

499   Linton, D. K., & Power, J. L. (2013). The personality traits of workplace bullies are often shared

500   by their victims: Is there a dark side to victims. *Personality and Individual Differences*, *54*,

501   738-743.

502   Lucas, R. E., & Donnellan, M. B. (2013). Improving the replicability and reproducibility of

503   research published in the Journal of Research in Personality. *Journal of Research in*

504   *Personality, 4*(47), 453-454.

505   Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality

506   assumption in large public health data sets. *Annual Review of Public Health*, *23*(1), 151–

507   169.

508   Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research how

509   often do they really occur? *Perspectives on Psychological Science, 7*(6), 537-542.

510   Morizot, J. M., Ainsworth, A. T., & Reise, S. P. (2007). Towards modern psychometrics:

511   Application of item response theory models in personality research. In R. W. Robins, R. C.

512   Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology*

513   (pp. 407–423). New York, NY: Guilford Press.

514   Olaru, G., Witthöft, M., & Wilhelm, O. (2015). Methods Matter: Testing Competing Models for

515   Designing Short-scale Big-Five Assessments. *Journal of Research in Personality*, *59*, 56-

516   68.

517   Paulhus, D. L. & Williams, K. M. (2002). The Dark Triad of personality: narcissism,

518   Machiavellianism, and psychopathy. *Journal of Research in Personality, 36*, 556–563.

519   R Core Team (2015). *R: A language and environment for statistical computing [computer*

520   *software*]. R Foundation for Statistical Computing, Vienna, Austria.

521   Raskin, R. N., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports,*

522   *45*(2), 590-590.

523   Reise, S. P., & Waller, N. G. (2009). Item response theory and clinical measurement. *Annual*

524   *Review of Clinical Psychology*, *5*, 27-48.

525   Revelle, W., & Wilt, J. (2013). The general factor of personality: A general critique. *Journal of*

526   *Research in Personality*, *47*(5), 493-504.

527   Schimmack, U. (2012). The ironic effect of significant results on the credibility of multiple-study

528   articles. *Psychological Methods*, *17*(4), 551.

529   Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected

530   in the social sciences. *Review of General Psychology*, *13*(2), 90–100.

531   Twenge, J. M., Campbell, W. K., & Freeman, E. C. (2012). Generational differences in young

532   adults' life goals, concern for others, and civic orientation, 1966–2009. *Journal of*

533   *Personality and Social Psychology, 102*, 1045-1062.

534   van der Linden, S., & Rosenthal, S. A. (2015). Measuring narcissism with a single question? A

535   replication and extension of the Single-Item Narcissism Scale (SINS). Personality and

536   Individual Differences, 90, 238-241.

537   Webster, G. D., & Jonason, P. K. (2013). Putting the "IRT" in "Dirty": Item response theory

538   analyses of the Dark Triad Dirty Dozen—An efficient measure of narcissism, psychopathy,

539   and Machiavellianism. *Personality and Individual Differences*, *54*(2), 302-306.

540   Zuckerman, M. (1989). Personality in the third dimension: A psychobiological approach.

541   *Personality and Individual Differences*, *10*(4), 391-418.

542   Zuckerman, M. (1991). *Psychobiology of personality*. Cambridge: Cambridge University Press.

543   Zuckerman, M., Kuhlman, D. M., Thornquist, M., & Kiers, H. (1991). Five (or three) robust

544   questionnaire scale factors of personality without culture. *Personality and Individual*

545   *Differences*, *12*(9), 929-941.

546   Table 1

547   *The Dark Triad Dirty Dozen Scale' traits, item numbers, statements and the key word in each*

548   *one of the statements.*

Trait	Item No.	Statement	Key Word
MACHIAVELLIANISM	1	I tend to manipulate others to get my way.	Manipulate
	2	I have used deceit or lied to get my way.	Deceit
	3	I have use flattery to get my way.	Flatter
	4	I tend to exploit others towards my own end.	Exploit
PSYCHOPATHY	5	I tend to lack remorse.	Remorse
	6	I tend to be unconcerned with the morality of my actions.	Amoral
	7	I tend to be callous or insensitive.	Callous
	8	I tend to be cynical.	Cynical
NARCISSISM	9	I tend to want others to admire me.	Admire
	10	I tend to want others to pay attention to me.	Attend
	11	I tend to seek prestige or status.	Status
	12	I tend to expect special favors from others.	Favors

549   *Note.* From Jonason, P. K., & Webster, G. D. (2010). The Dirty Dozen: A concise measure of the

550   Dark Triad. *Psychological Assessment*, *22*, 420-432.

551   Table 2

552   *Descriptive Analysis of the Dark Triad traits as measured by the Dark Triad Dirty Dozen.*

	M	SD	α	Skewness	Kurtosis	M_{male}	SD_{male}	$α_{male}$	M_{female}	SD_{female}	$α_{female}$
1 DD Machiavellianism	3.00	1.41	0.80	0.38	-0.66	3.23	1.45	.84	2.79	1.35	.85
2 DD Psychopathy	2.42	1.26	0.76	0.95	0.43	2.74	1.30	.81	2.13	1.14	.79
3 DD Narcissism	3.55	1.44	0.81	-0.15	-0.81	3.71	1.43	.75	3.41	1.44	.74
4 Dark Triad[+]	2.99	1.08	0.85	0.26	-0.27	3.23	1.08	.80	2.78	1.04	.81

553   *Note.* $N = 3,698$; $N_{males} = 1726$; $N_{females} = 1972$; [+] composite score of the three dark traits.

554

555   Table 3

556   *Convergent analysis (Persons' r) of the Dark Triad Dirty Dozen traits and the dark traits and Extraversion, and Neuroticism as*

557   *measured by the Mach-IV, Eysenck's Personality Questionnaire Revised, and the Neuroticism Personality Inventory.*

	1	2	3	4	5	6	7	8	9	**Item 4:**
1 DD Machiavellianism	—	.58	.50	.88	.49	.27	.37	.26	.10	.80
2 DD Psychopathy		—	.27	.75	.57	.41	.30	.25	-.11	.60
3 DD Narcissism			—	.76	.22	.09	.47	.20	.25	.43
4 Dark Triad [+]				—	.53	.31	.49	.30	.11	.77
5 Machiavellianism					—	.40	.34	.28	-.06	.50
6 Psychopathy						—	.35	.05	.05	.30
7 Narcissism							—	-.07	.40	.40
8 Neuroticism								—	-.24	.18
9 Extraversion									—	.09
Item 4: Exploit										—

558   *Note.* $N = 500$. All *r coefficients* > .12 are significant at $p < .01$. [+] Summarized composite score of the three dark traits.
559   Yellow fields: intra-relationships within the Dark Triad Dirty Dozen traits and the Dark Triad composite.
560   Blue fields: intra-relationships within the Dark Triad traits measured by the single instruments (i.e., MACH-IV, Eysenck's Personality
561   Questionnaire Revised, and the Narcissism Personality Inventory).
562   Green fields: relationships between corresponding dark traits measured using the Dirty Dozen scale and the single instruments.
563   Black fields: relationships between dark traits as measured by the Dirty Dozen and Neuroticism and Extraversion.
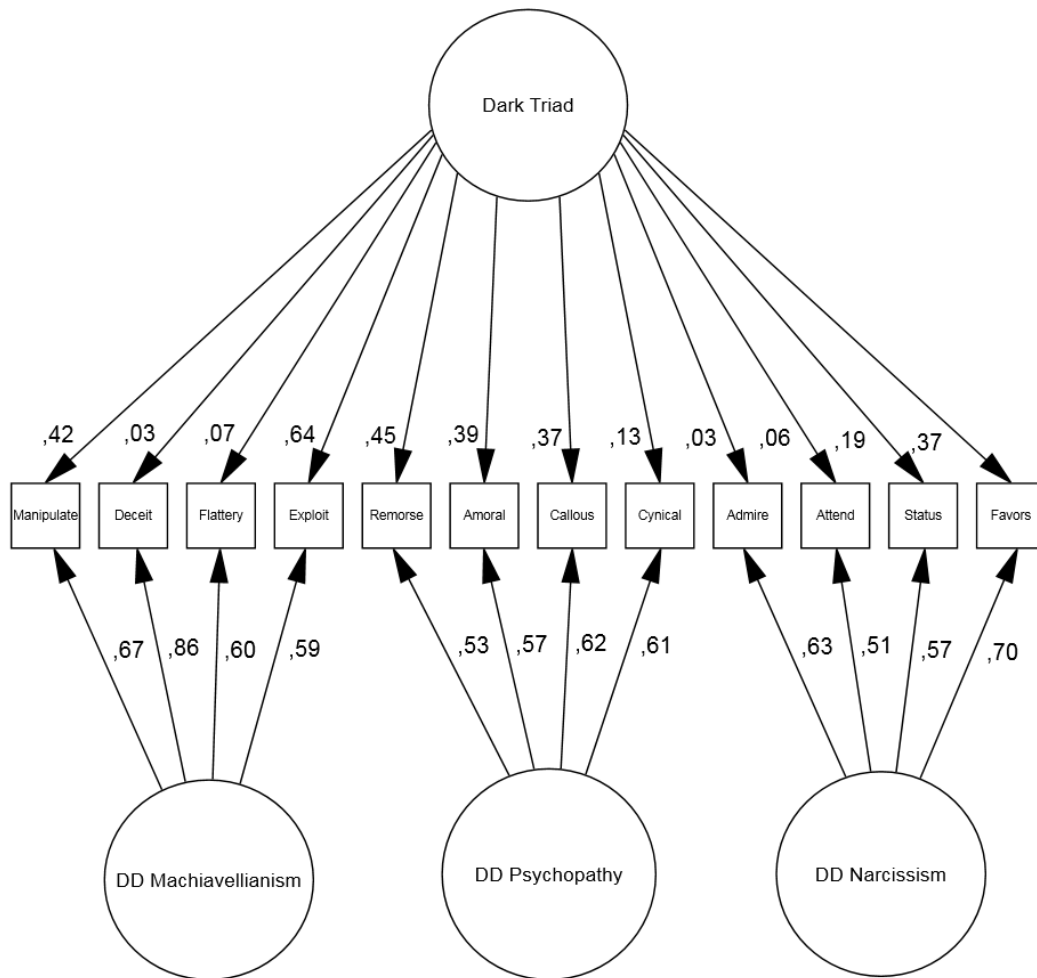564   Grey fields: relationships between dark traits as measured by the single instruments and Neuroticism and Extraversion.

565   Table 4
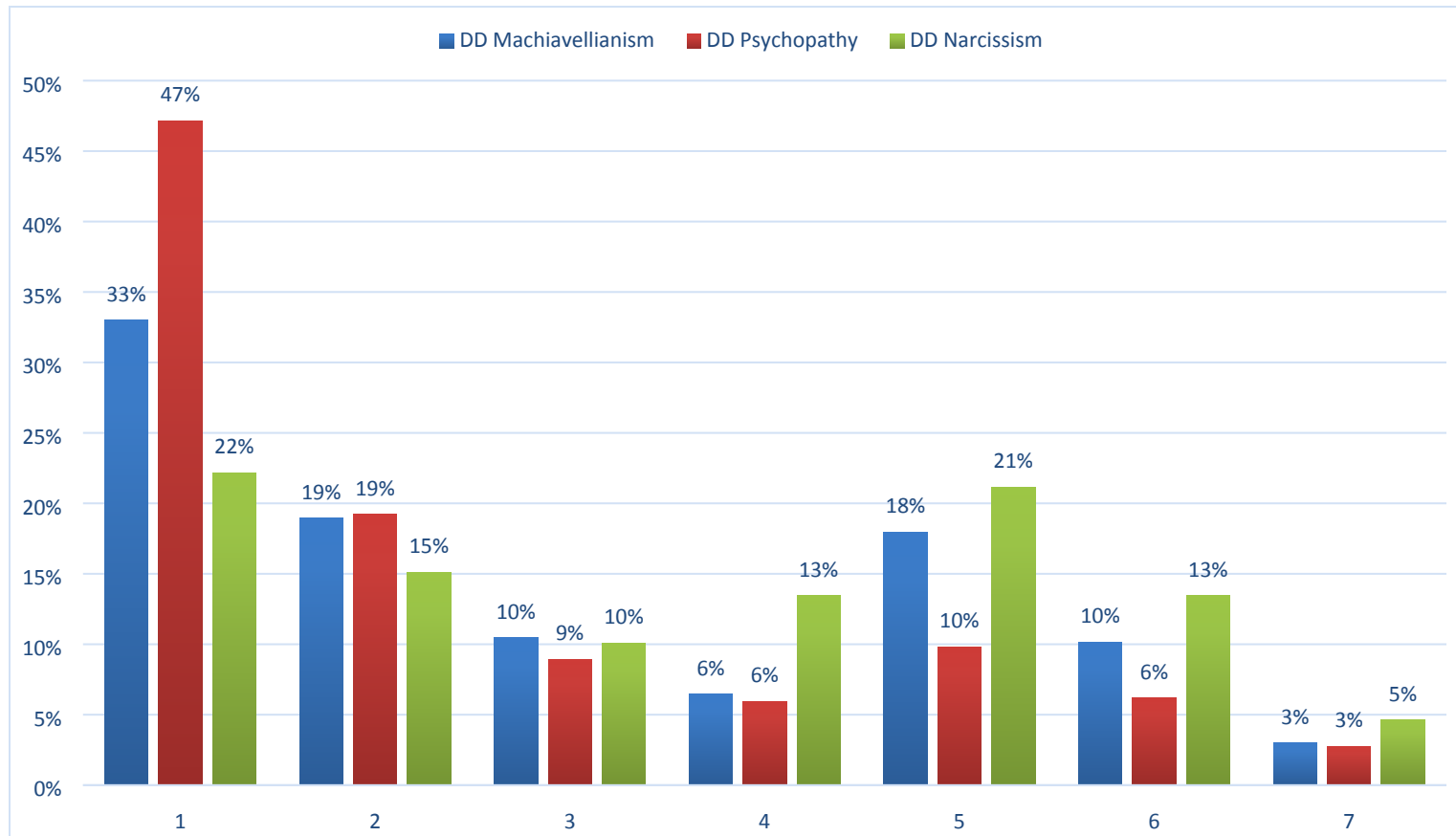
566   *Item Response Theory Analysis of the Dirty Dozen*

Item	a	b_1	b_2	b_3	b_4	b_5	b_6
4 Exploit	3.33	-0.03	0.63	0.97	1.31	1.99	2.58
1 Manipulate	2.73	-0.24	0.42	0.75	0.94	1.78	2.54
2 Deceit	1.96	-0.71	-0.07	0.27	0.44	1.24	2.35
6 Amoral	1.91	0.36	1.11	1.56	1.87	2.44	3.14
5 Remorse	1.84	0.25	1.00	1.36	1.68	2.29	3.10
7 Callous	1.82	-0.03	0.75	1.15	1.45	2.18	3.05
12 Favors	1.71	-0.47	0.42	0.82	1.37	2.21	3.35
3 Flatter	1.41	-1.43	-0.79	-0.35	-0.07	1.00	2.46
11 Status	1.19	-1.23	-0.45	0.02	0.52	1.59	2.76
8 Cynical	1.07	-1.13	-0.41	0.05	0.43	1.41	2.77
9 Admire	1.02	-2.12	-1.38	-0.89	-0.18	1.07	2.83
10 Attend	0.91	-1.99	-1.12	-0.54	0.21	1.72	3.75

567   *Note.* Items are ranked according to item's ability to discriminate (a) levels of the latent trait (the
568   core of the Dark Triad) and are numbered according to their positions in the original
569   questionnaire (DD Machiavellianism, 1-4, DD psychopathy, 5-8, and DD narcissism, 9-12). b_{1-6}
570   reports the item difficulties, reflecting the threshold level (-3 to +3 SD) of the latent trait
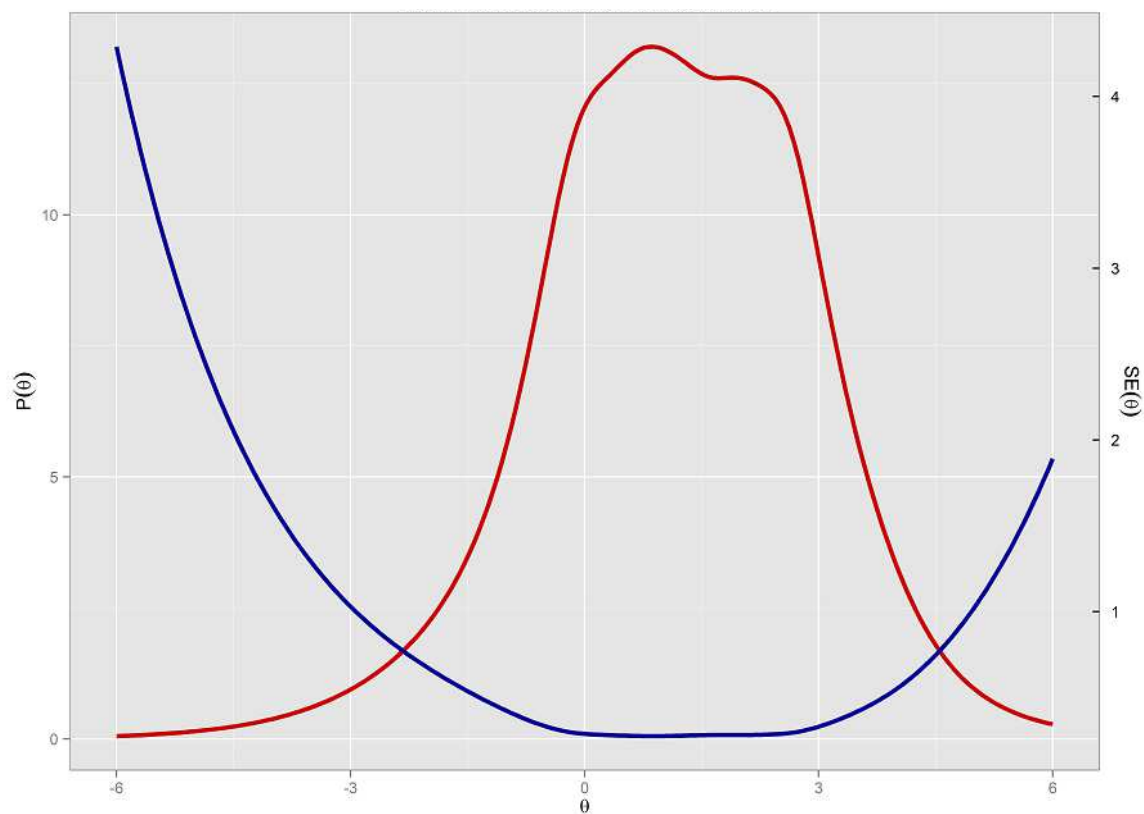571   necessary to have at least a 50% chance of endorsing the next scale-steps.
572

573
574   *Figure 1*. Bi-factor model of Dirty Dozen. *N* = 3,698. NFI = .98, CFI = .98, RMSEA = .05.

575

Figure 2. Frequency distributions showing the bi-modality of the three Dirty Dozen Dark Triad traits. $N = 3,698$. The numbers on the

y-axis represent the proportion of replies for each Likert-category (1–7) on the x-axis. For instance, 47% of total replies on DD

psychopathy items were placed on the lowest option (1) "strongly disagree", which depicts the skewness in response pattern.

579

580   *Figure 3*. Total Information Curve on the latent core of the Dark Triad Dirty Dozen. Overall, the

581   Dirty Dozen instrument delivers the most predictive value with persons with average and high

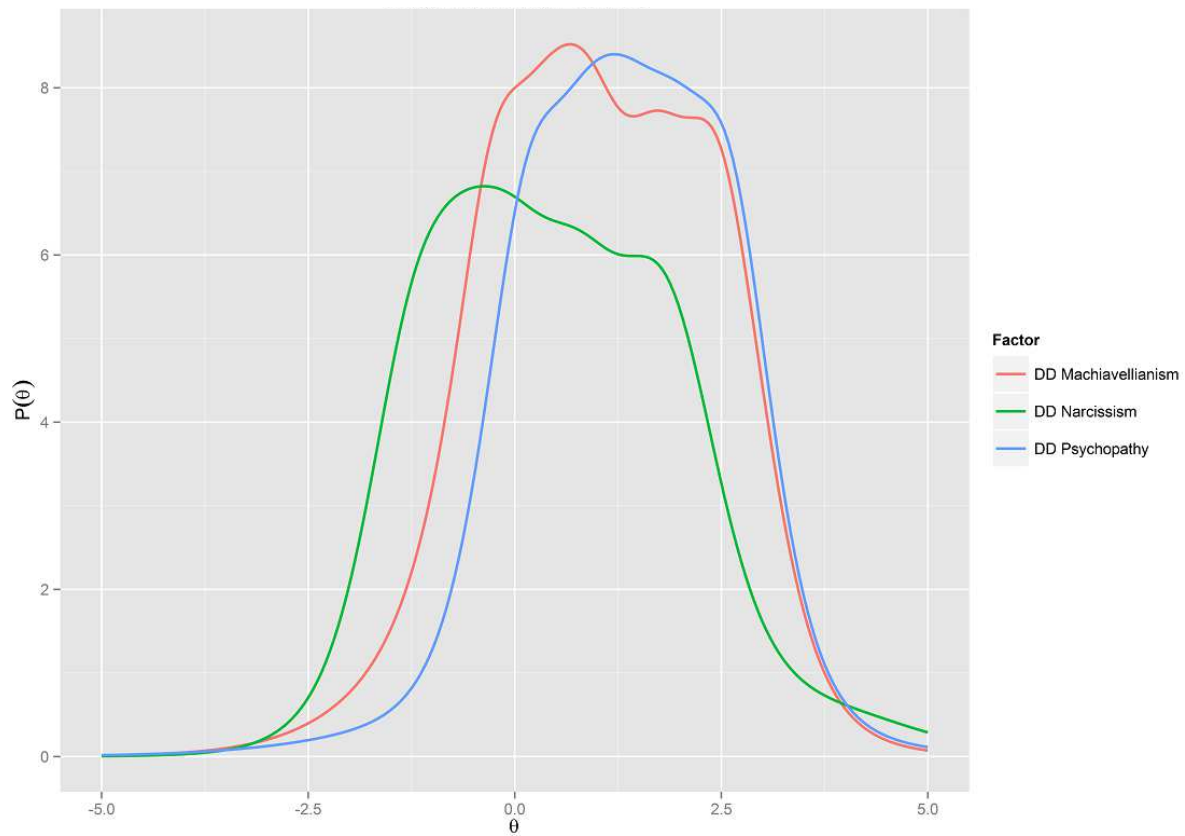582   Dark Triad traits (Theta > -0.5).

583

584

585

586

587

588

589

590

591

592

593

594
595   *Figure 4.* Scale Information Curves depicting the information content in each respective sub

596   factor.

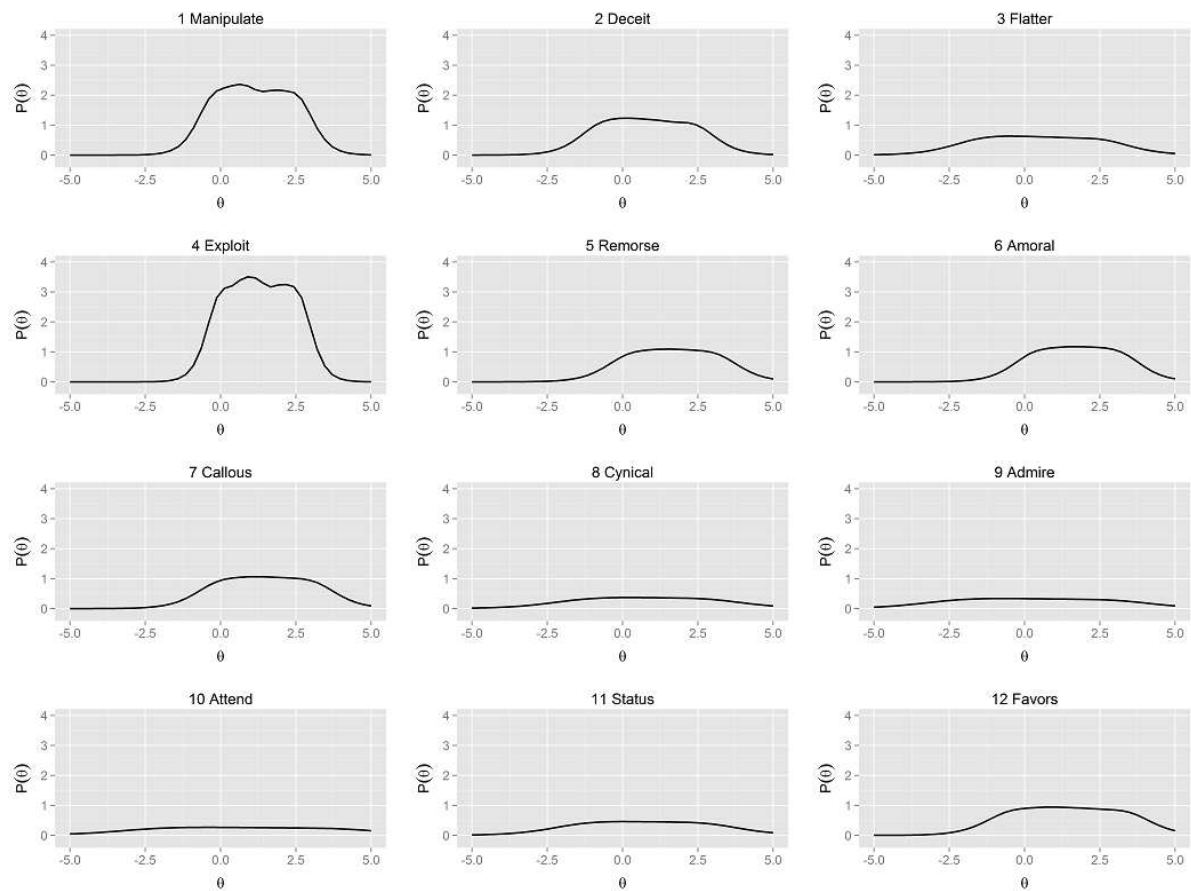597

598

599

600

601

602

603

604

605

606

607

608

609   *Figure 5*. Item Information Curves for each of the 12 items of the Dirty Dozen. Note that item 1

610   (Manipulate) and 4 (Exploit) deliver the most information on the latent Dark Triad (labeled

611   Theta). Item 8 (Cynicism) delivers the least information. Furthermore, the three first items on

612   narcissism (9-11) don't deliver much predictive information (flat curves, cf. earlier CFA bi-

613   factor model) to the overall Dark Triad trait – However, an IRT with only the narcissism items

614   confirms that these predict the latent trait (Theta, i.e., narcissism) very satisfactorily (cf. earlier

615   CFA hierarchical model).

616

617

618

619

620

621

622

623 *SUPPLEMENTAL MATERIAL*

624 The (mis)measurement of the Dark Triad Dirty Dozen: Exploitation at the core of the scale

625 Petri J. Kajonius[1, 2, 3], Björn N. Persson[3, 4], Patricia Rosenberg[3], Danilo Garcia[1, 3, 5, 6]*

626 [1]Department of Psychology, University of Gothenburg, Sweden

627 [2]Department of Social Psychology, University of Skövde, Sweden

628 [3]Network for Empowerment and Well-Being, Gothenburg, Sweden

629 [4]Department of Cognitive Neuroscience and Philosophy, University of Skövde, Sweden

630 [5]Blekinge Center of Competence, Blekinge County Council, Karlskrona, Sweden

631 [6]Centre for Ethics, Law, and Mental Health (CELAM), University of Gothenburg, Sweden

632

633

634

635

636

637   * Correspondence concerning this article should be addressed to D. Garcia, Network for

638   Empowerment and Well-Being, Axel W. Andersson 8A, SE 371 62 Lyckeby, Sweden. E-mail:

639   danilo.garcia@icloud.com.

640   Table S1

641   *Item Response Theory Rank and Exploratory Factor Analysis of Dirty Dozen Items.*

	Info rank	Males			Females		
		Machiavellianism	Psychopathy	Narcissism	Machiavellianism	Psychopathy	Narcissism
1 Manipulate*	2	**.81**	*.54*	.34	**.77**	.50	.38
2 Deceit	3	**.73**	.46	.30	**.76**	.40	.35
3 Flatter	8	**.59**	.27	.44	**.60**	.26	.43
4 Exploit*	1	**.78**	*.60*	.36	**.70**	*.60*	.35
5 Remorse	5	.44	**.78**	.08	.42	**.84**	.17
6 Amoral	4	.46	**.72**	.15	.41	**.76**	.16
7 Callous*	6	.47	**.77**	.13	.50	**.70**	.20
<u>8 Cynical*</u>	11	.36	**<u>.40</u>**	.14	.45	**<u>.42</u>**	.25
9 Admire	10	.35	.09	**.81**	.40	.11	**.81**
10 Attend	12	.33	.08	**.78**	.34	.11	**.80**
11 Status	9	.38	.16	**.73**	.41	.23	**.73**
12 Favors*	*7*	*.52*	.36	**.55**	*.51*	.43	**.57**

642   *Note.* * = Potentially conflicting items in regards to double-loadings. Bold figures = Loadings

643   according to three-factor Dark Triad theory. Cursive figures = Loadings deviating from three-

644   factor Dark Triad theory (> .50). Item 8, underscored, "I tend to be cynical" does not load well

645   enough on the psychopathy-factor and deviates from theory by aligning also with

646   Machiavellianism (especially for females). Items 1, "I tend to manipulate others to get my way",

647   4, "I tend to exploit others toward my own end", and 12, "I tend to expect special favors from

648   others" double-load on two separate factors. Items 4, 5, and 7 have somewhatvarying loadings

649   for males and females.