

The Dirty Dozen: A Concise Measure of the Dark Triad

Peter K. Jonason
University of West Florida

Gregory D. Webster
University of Florida

There has been an exponential increase of interest in the dark side of human nature during the last decade. To better understand this dark side, the authors developed and validated a concise, 12-item measure of the Dark Triad: narcissism, psychopathy, Machiavellianism. In 4 studies involving 1,085 participants, they examined its structural reliability, convergent and discriminant validity (Studies 1, 2, and 4), and test–retest reliability (Study 3). Their measure retained the flexibility needed to measure these 3 independent-yet-related constructs while improving its efficiency by reducing its item count by 87% (from 91 to 12 items). The measure retained its core of disagreeableness, short-term mating, and aggressiveness. They call this measure the *Dirty Dozen*, but it cleanly measures the Dark Triad.

Keywords: Dark Triad, narcissism, Machiavellianism, psychopathy, measurement

The Dark Triad is a term used to describe a constellation of three socially undesirable personality traits: narcissism, psychopathy, and Machiavellianism (Paulhus & Williams, 2002). Research on the Dark Triad has increased exponentially over the last decade. An analysis of Google Scholar hit counts for “Dark Triad” in scientific works reveals an explosive increase from one in 2002 to at least 38 in 2009. Despite the recent flurry of scientific interest in the Dark Triad, it has a substantial methodological shortcoming: With over 90 items spread across three scales, it is remarkably inefficient for researchers to measure. In this article, we propose a much-needed solution to this problem: We develop and test the psychometric properties of an efficient, 12-item version of the Dark Triad called the *Dirty Dozen*.

There are at least two reasons to develop a concise measure of the Dark Triad. First, each measure has its own response biases and limitations. For instance, the Mach IV (Christie & Geis, 1970), a measure of Machiavellianism, may be biased by social desirability (Wilson, Near, & Miller, 1996) and has some unclear psychometric properties (Hunter, Gerbing, & Boster, 1982; for an exception, see Jones & Paulhus, 2009). Indeed, across two recent studies (Jonason, Li, Buss, 2010; Jonason, Li, & Teicher, in press), the internal consistency of the Mach IV did not exceed .70, a remarkably low number for a 20-item scale (Carmines & Zeller, 1979). The Narcissistic Personality Inventory (NPI; Raskin & Terry, 1988), a measure of narcissism, is subject to high rates of impression management (Auerbach, 1984) and is composed of a series of dichotomous questions, which can be problematic (Comrey, 1973). In addition, the use of two different measurement techniques further complicates one’s ability to measure the Dark Triad, re-

quiring scores on each measure to be standardized (Jonason, Li, Webster, & Schmitt, 2009).

Second, assessing the Dark Triad’s 91 items is inefficient, time-consuming, and may cause response fatigue in some participants. When studying the Dark Triad and one or more other measures of interest (e.g., self-esteem, Big Five personality traits), the total number of items in a questionnaire can easily exceed 100. Large-scale surveys have the advantage of providing a plethora of data, but it may come at the cost of response errors resulting from participant fatigue. Concise measures can eliminate redundant items, save time and effort, and thus reduce participant fatigue and frustration (Saucier, 1994). Using measures that are more efficient can therefore be mutually beneficial to both researchers and participants in terms of time and resources saved, all without sacrificing precision, so long as the concise measure adequately reflects its original version.

Traditionally, the Dark Triad is composed of three independent constructs with some overlap (e.g., Paulhus & Williams, 2002). However, recent evidence suggests there are good theoretical and empirical reasons to treat them as different measures of the same latent construct. Specifically, the Dark Triad as a whole can be thought of as a short-term, agentic, exploitive social strategy that may have evolved to enable exploitation when conspecifics are likely to avoid or punish defectors (Jonason et al., 2009; Jonason et al., in press). Empirical evidence suggests that narcissism, psychopathy, and Machiavellianism measure a single, latent construct that accounts for approximately 50% of the variance associated with the three scales (Jonason, Li, & Buss, 2010; Jonason et al., 2009). Moreover, the Dark Triad’s correlations with mating motivations (Jonason et al., 2009) and agreeableness (Jakobwitz & Egan, 2006; Paulhus & Williams, 2002) are similar across all three dimensions and sometimes stronger with a single, latent composite (Jonason et al., 2009; Jonason et al., in press). Thus, it is important that any concise measure of the Dark Triad incorporate the flexibility of being scored as either three related subscales or as a single, composite scale.

In addition to retaining flexibility, the Dark Triad Dirty Dozen should behave in ways that the longer measures do. First, the Dirty Dozen should be correlated with the longer, original measures of

Peter K. Jonason, Department of Psychology, University of West Florida; Gregory D. Webster, Department of Psychology, University of Florida.

The authors thank Jeremy Tost, Kayla Whitworth, Tiffany Rodriguez, Joshua Legarreta, Austin Lee Nichols, and Jessica Darby for help with data collection and entry.

Correspondence concerning this article should be addressed to Peter K. Jonason, University of West Florida, Department of Psychology, Bldg. 41, 11000 University Parkway, Pensacola, FL 32514. E-mail: peterkarljonason@yahoo.com

the Dark Triad. Second, the Dirty Dozen should be correlated negatively with agreeableness (e.g., Paulhus & Williams, 2002) and positively with a short-term mating strategy (Jonason et al., 2009) and aggressiveness (e.g., Bushman & Baumeister, 1998; Paulhus & Williams, 2002). It is common to use the Big Five and other “normal,” lower order personality traits to describe the validity of measures (e.g., Schutte et al., 1998; Seemann, Buboltz, Thomas, Soper, & Wilkinson, 2005). Last, there is ample evidence that suggests that men score higher on all three of these traits than women do (e.g., Jonason, Li, & Buss, 2010; Jonason et al., 2009), and therefore, we expect that men should score higher than women do on the Dirty Dozen measures. These predictions constitute validity tests of our measure (Cronbach & Meehl, 1955).

In the present research, we sought to develop a concise measure of the Dark Triad that improves its efficiency by reducing its item count by 87% (from 91 to 12 items), while simultaneously preserving its flexibility in serving as either a one- or three-dimensional construct. In two studies, we develop this measure through principal components analyses (PCAs) and confirmatory factor analyses (CFAs). We also validate the Dirty Dozen through assessment of the surrounding nomological network (Cronbach & Meehl, 1955) with constructs that have proven important in prior research such as the original, 91-item version of the Dark Triad, the Big Five, mating, self-esteem level and stability, and aggression. In a third study, we assess the test–retest reliability of the measures over a 3-week period. Consistency over time is one of the defining features of personality traits, and therefore, such evidence will bolster our claims about the usefulness of our measure as well as provide support for the treatment of the Dark Triad as personality variables. In a fourth study, we fine tune our measure by simplifying a double-barreled item, improve the internal consistency of the scale, and again confirm the Dark Triad Dirty Dozen’s factor structure.

Study 1

In Study 1, we developed a new, concise, and psychometrically sound measure of the Dark Triad. We first created 22 candidate items inspired by the original Dark Triad measures that we felt were the most theoretically central to each construct. We tested our measure by correlating it with the original Dark Triad measures, the Big Five, and measures of mating. We also provide evidence for sex differences among these measures.

Method

Participants and procedures. Two hundred seventy-three psychology students (90 men, 183 women) aged 18–47 years ($M = 20.08$, $SD = 3.79$) from the Southwestern United States received course credit for completing the surveys described below. Participants completed packets in a lab setup for mass testing where as many as 10 other people could participate at a time. Participants were instructed to ensure that at least one seat separated them from other participants. Once they completed the measures, they were debriefed and thanked for their participation.

Measures. To assess the Big Five personality dimensions, we used the Big Five Inventory (Benet-Martinez & John, 1998), a cross-culturally validated instrument, using a response scale from 1 (*strongly disagree*) to 5 (*strongly agree*). Five factors were

detected: Extraversion (Cronbach’s $\alpha = .84$, 8 items), Neuroticism ($\alpha = .79$, 8 items), Openness ($\alpha = .76$, 10 items), Conscientiousness ($\alpha = .69$, 9 items), and Agreeableness ($\alpha = .72$, 9 items).

Narcissism was assessed with the 40-item NPI, a validated and widely used measure (Raskin & Terry, 1988). For each item, participants chose one of two statements they felt applied to them more. One of the two statements reflected a narcissistic attitude (e.g., “I have a natural talent for influencing people”), whereas the other statement did not (e.g., “I am not good at influencing people”). We summed the total number of narcissistic statements the participants endorsed as an index of narcissism ($\alpha = .80$).

The 31-item Self-Report Psychopathy Scale–III (Paulhus, Hemphill, & Hare, in press) assessed subclinical psychopathy. Participants rated how much they agreed (1 = *strongly disagree*, 5 = *strongly agree*) with statements such as “I enjoy driving at high speeds” and “I think I could beat a lie detector.” Items were averaged to create an index of psychopathy ($\alpha = .74$).

Machiavellianism was measured with the 20-item Mach IV (Christie & Geis, 1970). Participants were asked how much they agreed (1 = *strongly disagree*, 5 = *strongly agree*) with statements such as “It is hard to get ahead without cutting corners here and there” and “People suffering from incurable diseases should have the choice of being put painlessly to death.” The items were averaged to create a Machiavellianism index ($\alpha = .65$).

The three Dark Triad measures can be treated as one measure (Jonason et al., 2009). We standardized (z -scored) the overall scores on each of the three scales and then averaged all three standardized scores together to create a composite Dark Triad score.

Sociosexuality was measured as a tripartite personality construct (Jackson & Kirkpatrick, 2007). We replicated all three dimensions. The items for each dimension were averaged to create a measure of short-term mating orientation ($\alpha = .94$), long-term mating orientation ($\alpha = .91$), and sexual experiences ($\alpha = .75$).

Results and Discussion

Factor structure. We conducted separate PCAs and internal consistency analyses for each measure (see Table 1). All PCAs used oblique rotation, and loadings were from pattern matrices. Eigenvalues greater than one were used to determine factors. The four items with the strongest loadings on the primary factor were chosen from each of the three Dark Triad measures. Together, these 12 items constituted the Dark Triad Dirty Dozen. Using the same methods, we then conducted a single PCA and internal consistency analysis on the Dirty Dozen to test the factor structure. As predicted, three factors emerged: Machiavellianism, Psychopathy, and Narcissism (see Table 2). The correlations among these rotated factors were modest ($|r|s \leq .35$; see Table 3). Overall, the Dark Triad Dirty Dozen achieved good internal consistency ($\alpha = .83$). When evaluated separately, the internal consistency (α) for each component improved after being reduced to four items for both psychopathy (from .62 to .63) and Machiavellianism (from .67 to .72). In contrast, narcissism’s internal consistency decreased (from .87 to .79), which is not surprising considering that α increases with the number of items (all else being equal) and that narcissism had the largest decline in items (from 11 to four).

Along the same lines, although the α of .63 for psychopathy may seem low (Nunnally, 1978), it is respectable for a scale with

Table 1

Separate Subscale-Based Principal Components Analysis Using Oblique Rotation of—and Item–Scale Correlations for—a 22-Item Dark Triad Measure in Two Studies

Item	Study 1 (N = 273)			Study 2 (N = 246)			
	Factor			Factor			
	1	2	<i>r</i>	1	2	3	<i>r</i>
Narcissism (α s = .87 and .85)							
I tend to want others to admire me.	.90	.12	<i>.64</i>	.84	–.03	–.00	<i>.63</i>
I tend to want others to pay attention to me.	.82	.07	<i>.61</i>	.84	–.06	–.05	<i>.60</i>
I tend to expect special favors from others.	.71	.06	<i>.50</i>	.54	–.53	.13	<i>.51</i>
I tend to seek prestige or status.	.62	–.21	<i>.66</i>	.74	.10	.02	<i>.55</i>
I tend to feel that things are owed to me.	.58	–.15	<i>.56</i>	.26	–.59	.32	<i>.43</i>
I tend to try to be dominant in social situations.	.57	–.11	<i>.52</i>	.32	.00	.39	<i>.51</i>
I tend to be grandiose or pompous.	.34	–.34	<i>.50</i>	–.25	–.10	.84	<i>.40</i>
I tend to feel that I am more special than others.	–.06	–.93	<i>.64</i>	.15	.30	.70	<i>.60</i>
I tend to feel that I am better than others.	–.05	–.90	<i>.62</i>	.22	.11	.72	<i>.70</i>
I tend to have a sense of self-importance.	.06	–.72	<i>.57</i>	.35	.64	.30	<i>.42</i>
I tend to be egocentric.	.27	–.42	<i>.51</i>	.05	–.11	.66	<i>.50</i>
Psychopathy (α s = .62 and .66)							
I tend to lack remorse.	.81	–.04	<i>.40</i>	.82	–.00	—	<i>.50</i>
I tend to be callous or insensitive.	.74	.16	<i>.48</i>	.83	–.05	—	<i>.47</i>
I tend to not be too concerned with morality or the morality of my actions.	.65	–.06	<i>.31</i>	.53	–.04	—	<i>.26</i>
I tend to be cynical.	.55	–.01	<i>.27</i>	.67	.12	—	<i>.44</i>
I tend to get frustrated easily.	–.03	.88	<i>.34</i>	–.02	.89	—	<i>.34</i>
I tend to lose my temper quickly.	–.04	.85	<i>.36</i>	.02	.88	—	<i>.35</i>
Machiavellianism (α s = .67 and .72)							
I have used deceit or lied to get my way.	.76	—	<i>.55</i>	.81	—	—	<i>.62</i>
I tend to manipulate others to get my way.	.76	—	<i>.50</i>	.77	—	—	<i>.57</i>
I have used flattery to get my way.	.72	—	<i>.47</i>	.75	—	—	<i>.54</i>
I tend to exploit others towards my own end.	.69	—	<i>.46</i>	.69	—	—	<i>.48</i>
I tend to have trouble understanding other people's feelings.	.34	—	<i>.20</i>	.38	—	—	<i>.24</i>

Note. Factor loadings $\geq |.40|$ and the four best items from each subscale (i.e., the Dirty Dozen items) are boldfaced. *r* = corrected item–scale correlation (in italics).

only four items (see Carmines & Zeller, 1979). For example, for a typical scale that has a mean interitem correlation of .30, α is .63 for a four-item scale, but it increases to .81 for a 10-item scale, or .90 for a 20-item scale, solely on the basis of additional items. Thus, adjusting for number-of-item inflation, α s in the .60s are reasonable for four-item scales. We suspect that psychopathy's comparatively lower internal consistency may be the result of the double-barreled nature of the item, "I tend to not be too concerned with morality or the morality of my actions," which is an issue we address in Study 4.

Convergent and discriminant validity. Next, we assessed the convergent and discriminant validity of the Dirty Dozen (see Table 4). We first examined the relationships among our concise measures and the original versions of the Dark Triad and its three components using a multitrait–multimethod matrix (Campbell & Fiske, 1959). The top of Table 4 shows the heteromethod block of interest. As expected, the heteromethod block of the multitrait–multimethod matrix revealed a consistent pattern of convergent and discriminant validity among the Dark Triad components, with the on-diagonal correlations (validity diagonals) being stronger than the off-diagonal correlations (heterotrait–heteromethod triangles) with two exceptions. First, the 31-item measure of Psychop-

athy correlated slightly more strongly with our concise measure of Machiavellianism (.44) than it did with our concise measure of psychopathy (.42). Second, the 31-item measure of Psychopathy also correlated slightly more strongly with the Dirty Dozen (.51) than the Dirty Dozen did with the original 91-item Dark Triad composite measure (.47). Despite these exceptions, the multitrait–multimethod matrix largely supported the expected pattern of convergent and discriminant validity between the original 91-item Dark Triad and our 12-item Dirty Dozen.

Assessments of the nomological network were also telling. The Dirty Dozen measures retained a core of disagreeableness (Paulhus & Williams, 2002), showed low levels of conscientiousness (Jonason et al., in press), and were more closely correlated with a short-term mating orientation than a long-term one (Jonason et al., 2009). This suggests that the Dirty Dozen taps the same personality traits as the unabridged Dark Triad measures.

Sex differences. Tests for sex differences revealed men scored higher than women on our concise scales of Machiavellianism, $t(265) = 3.98, p < .01, d = 0.49$; narcissism, $t(265) = 3.11, p < .01, d = 0.40$; and psychopathy, $t(265) = 4.95, p < .01, d = 0.62$; as well as the Dirty Dozen, $t(265) = 4.95, p < .01, d = 0.64$.

Table 2
Principal Components Analysis Using Oblique Rotation of and Item–Scale Correlations for the Dirty Dozen Dark Triad Items in Studies 1 and 2 and Item-Level Temporal Reliability in Study 3

Item	Study 1 (N = 273)				Study 2 (N = 246)				Study 3 (N = 60)
	Factor				Factor				
	M	P	N	<i>r</i>	M	P	N	<i>r</i>	Item-level α across 3 time points
1. I tend to manipulate others to get my way.	.83	-.01	-.04	<i>.50</i>	.71	.20	.06	<i>.54</i>	.87
2. I have used deceit or lied to get my way.	.62	.01	.18	<i>.54</i>	.82	.03	.03	<i>.54</i>	.91
3. I have use flattery to get my way.	.57	-.07	.30	<i>.52</i>	.77	-.05	-.04	<i>.49</i>	.94
4. I tend to exploit others towards my own end.	.54	.24	.15	<i>.58</i>	.54	-.02	-.34	<i>.55</i>	.93
5. I tend to lack remorse.	.08	.76	-.00	<i>.44</i>	.05	.76	-.08	<i>.52</i>	.79
6. I tend to not be too concerned with morality or the morality of my actions. ^a	-.25	.75	.21	<i>.35</i>	.31	.30	-.01	<i>.36</i>	.72
7. I tend to be callous or insensitive.	.46	.55	-.07	<i>.54</i>	-.04	.88	.00	<i>.46</i>	.86
8. I tend to be cynical.	.25	.43	-.10	<i>.30</i>	.06	.69	.01	<i>.41</i>	.92
9. I tend to want others to admire me.	.07	-.10	.85	<i>.56</i>	-.02	.00	-.86	<i>.54</i>	.91
10. I tend to want others to pay attention to me.	.13	-.12	.81	<i>.55</i>	.13	-.06	-.78	<i>.55</i>	.91
11. I tend to seek prestige or status.	-.09	.23	.74	<i>.55</i>	-.21	.22	-.76	<i>.46</i>	.92
12. I tend to expect special favors from others.	.16	.17	.55	<i>.55</i>	.18	-.07	-.64	<i>.47</i>	.93

Note. Factor loadings $\geq |.30|$ are boldfaced. M = Machiavellianism ($\alpha = .72$ and $.77$); P = Psychopathy ($\alpha = .63$ and $.69$); N = Narcissism ($\alpha = .79$ and $.78$); *r* = corrected item–scale correlation (in italics). Dirty Dozen Dark Triad $\alpha = .83$ for Studies 1 and 2.

^a Because of its double-barreled nature and low reliability, Item 6 was simplified to read, “I tend to be unconcerned with the morality of my actions” in Study 4. We encourage researchers to use the items listed in Table 8.

Study 2

In Study 2, we provide additional evidence of the Dirty Dozen’s sound psychometric properties and its validity as a concise measure of constructs underlying the Dark Triad. We assess the validity of this measure by correlating it with alternative measures of the Big Five, mating, and a global measure of self-esteem. We again test sex differences to verify that men score higher than women do on these measures.

Method

Participants and procedures. Two hundred forty-six psychology students (101 men, 145 women) aged 18–42 years ($M = 20.69$, $SD = 3.76$) from the Southwestern United States received course credit for completing the surveys described below. Procedures from Study 1 were replicated here.

Measures. To measure the Big Five, we used the Ten-Item Personality Inventory (TIPI; Gosling, Rentfrow, & Swann, 2003), which asks two questions for each dimension. Participants were asked, for instance, how much (1 = *not at all*, 5 =

very much) they think of themselves as “extraverted, enthusiastic” and “quiet, reserved” as measures of extraversion. Estimates of internal consistency returned low rates: extraversion ($\alpha = .55$), agreeableness ($\alpha = .22$), conscientiousness ($\alpha = .44$), neuroticism ($\alpha = .38$), and openness ($\alpha = .09$), as is to be expected for scales composed of a small number of items (Kline, 2000). Although these estimates are smaller than those reported in Gosling et al. (2003), such estimates are expected because internal consistency estimates are positively related to the number of scale items (Carmines & Zeller, 1979). Indeed, Gosling et al. (2003) made just such a point, noting that the more appropriate test of reliability for a brief measure is test–retest instead of Cronbach’s alpha (p. 516), evidence we provide in Study 3. However, because of the lower levels of internal consistency—even by liberal standards (Schmitt, 1996)—we corrected the correlations between the TIPI dimensions and the Dirty Dozen for attenuation from measurement error (see Cohen, Cohen, West, & Aiken, 2003).

Sociosexuality was measured using the seven-item Sociosexual Orientation Inventory (Simpson & Gangestad, 1991), which gauges participants’ attitudes and behaviors regarding sexual intercourse with multiple partners. For instance, participants were asked how much they agreed (1 = *strongly disagree*; 9 = *strongly agree*) with the statement “I can imagine myself being comfortable and enjoying casual sex with different partners.” Individual items were standardized (*z*-scored) prior to computing scale means and as averaged into an index ($\alpha = .80$).

Global self-esteem was measured with the 10-item Rosenberg’s (1965) Self-Esteem Scale. Participants were asked how much they agreed (1 = *strongly disagree*; 4 = *strongly agree*) with statements like: “I feel that I am person of worth, at least on an equal basis with others.” The 10 items were averaged to create an index of self-esteem ($\alpha = .80$).

Table 3
Principal Component Correlation Matrix for Study 1 (Above the Diagonal) and 2 (Below the Diagonal)

Variable	Machiavellianism	Psychopathy	Narcissism
Machiavellianism	—	.30 (.51**)	.34 (.55**)
Psychopathy	.35 (.51**)	—	.23 (.37**)
Narcissism	-.34 (.60**)	-.27 (.34**)	—

Note. Oblique rotation was used. The zero-order correlations among the unweighted composite measures are shown in parentheses for comparison. ** $p < .01$.

Table 4

Correlations Among the 12-Item Dark Triad Dirty Dozen, Its Components, and the 91-Item Dark Triad, Its Components, and Mating, Self-Esteem, Big Five Personality, and Aggression

Variable	Dark Triad Dirty Dozen			
	Machiavellianism	Psychopathy	Narcissism	Total
Study 1 (<i>N</i> = 273)				
Multitrait-multimethod matrix				
Machiavellianism (20-item Mach IV)	.34**	.29**	.17**	.33**
Psychopathy (31-item SRP-III)	.44**	.42**	.38**	.51**
Narcissism (40-item NPI)	.37**	.14*	.46**	.41**
Dark Triad composite	.44**	.37**	.34**	.47**
Mating				
Short-term mating orientation	.37**	.31**	.34**	.43**
Long-term mating orientation	-.12	-.17**	.06	-.08
Sexual experience	.23**	.12	.22**	.24**
Big Five personality traits				
Extraversion	.07	-.07	.15*	.06
Agreeableness	-.35**	-.45**	-.17**	-.24**
Conscientiousness	-.25**	-.19**	-.17**	-.24**
Neuroticism	-.03	-.07	-.10	-.08
Openness	-.03	.03	.15*	.07
Study 2 (<i>N</i> = 246)				
Self-esteem	-.09	-.09	-.13*	-.13*
Sociosexual Orientation Inventory	.22**	.29**	.21**	.31**
Ten-Item Personality Inventory (TIPI)				
Extraversion	.18** (.27**)	-.05 (-.08)	.10 (.15*)	.10 (.15*)
Agreeableness	-.24** (-.58**)	-.42** (-.69**)	-.17** (-.41**)	-.35** (-.82**)
Conscientiousness	-.20** (-.34**)	-.09 (-.16*)	-.14* (-.24**)	-.18** (-.30**)
Neuroticism	-.11 (-.20**)	-.14* (-.27**)	-.14* (-.27**)	-.17** (-.30**)
Openness	.04 (.15*)	-.00 (-.00)	.06 (.23**)	.05 (.18**)
Study 3 (<i>N</i> = 60)				
Self-esteem				
Single-item self-esteem scale	-.04	-.02	-.03	-.04
Labile self-esteem scale	.22	.25	.10	.24
Stability of self scale	.24	.35**	.16	.31*
Aggression Questionnaire (AQ)				
Physical aggression	.38** (.23)	.40** (.38**)	.13 (.04)	.39** (.26*)
Verbal aggression	.44** (.42**)	.40** (.38**)	.06 (.03)	.38** (.35**)
Hostility	.40** (.38**)	.23 (.22)	.44** (.43**)	.49** (.48**)
Anger	.22 (.25)	.21 (.22)	.12 (.12)	.23 (.26*)
Total	.50** (.43**)	.45** (.42**)	.24 (.20)	.51** (.45**)

Note. SRP-III = Self-Report Psychopathy Scale-III (Paulhus, Hemphill, & Hare, in press); NPI = Narcissistic Personality Inventory (Raskin & Terry, 1988). Convergent validity correlations are bolded. Parenthetic correlations are either corrected for attenuation from measurement error (TIPI) or partial correlations controlling for sex (AQ).

* $p < .05$. ** $p < .01$.

Results and Discussion

Factor structure. Using the same methods as in Study 1, we conducted separate PCAs and internal consistency analyses for each measure based on all 22 items (see Table 1). These results largely confirmed our findings from Study 1: The four items with the highest primary factor loadings for each component were the same. We then conducted a single PCA and internal consistency analysis on the Dirty Dozen to test its factor structure. As predicted, three factors emerged: Machiavellianism, Narcissism, and Psychopathy (see Table 2). The correlations among these rotated factors were modest ($|r|s \leq .35$; see Table 3). These findings replicated those of Study 1 with one minor exception: Item 6 (“I tend to not be too concerned with morality or the morality of my

actions”) loaded slightly higher on Machiavellianism (.31) than did its predicted dimension, Psychopathy (.30).

Each scale returned highly similar rates of internal consistency as reported in Study 1. When treated as a single scale, internal consistency analyses for the Dark Triad Dirty Dozen returned an identical Cronbach’s alpha ($\alpha = .83$) in Study 2 as in Study 1. Similar to Study 1, when we reduced the number of items to be four for each scale, our Machiavellianism (from .72 to .77) and Psychopathy (from .66 to .69) measures increased in internal consistency. In addition, like in Study 1, the internal consistency of our Narcissism measure decreased (from .85 to .78).

We ran a series of nested CFAs to evaluate our Dirty Dozen Dark Triad items. In the first model, we allowed all 12 items to

load on a single, one-dimensional, Dark Triad factor (see Figure 1A). The single-factor measurement model fit the data poorly, $\chi^2(54) = 344.50, p < .01, \chi^2/df = 6.38$; normed fit index (NFI) = .63, comparative fit index (CFI) = .66; root-mean-square error of approximation (RMSEA) = .15, 90% confidence interval (CI) = .13–.16, $p_{\text{close fit}} < .01$. Because they are structurally equivalent, the three-dimensional (see Figure 1B) and hierarchical (see Figure 1C) models produced identical fits to the data, which were reasonable, $\chi^2(51) = 123.90, p < .01, \chi^2/df = 2.43$; NFI = .87, CFI = .92; RMSEA = .08, 90% CI = .06–.09, $p_{\text{close fit}} < .01$. The multidimensional and hierarchical models fit the data significantly better than the one-dimensional model, $\Delta\chi^2(3) = 220.50, p < .01$.

Convergent and discriminant validity. Next, we assessed the convergent and discriminant validity of the Dirty Dozen (see Table 4). Again, the Dirty Dozen measures retained a core of disagreeableness (Paulhus & Williams, 2002) and were positively

associated with sociosexuality (Jonason et al., 2009; Webster & Bryan, 2007). When we corrected for attenuation for measurement error, we replicated evidence from above that the Dirty Dozen was negatively correlated with conscientiousness. We also found that scores on the Dirty Dozen and the TIPI's neuroticism dimensions were negatively correlated and scores on the Dirty Dozen and the TIPI's openness dimension were positively correlated.

Sex differences. Tests for sex differences revealed men scored higher than women on our concise scales of narcissism, $t(242) = 4.76, p < .01, d = 0.62$, and psychopathy, $t(242) = 3.18, p < .01, d = 0.42$, as well as the Dirty Dozen, $t(242) = 4.01, p < .01, d = 0.52$. Men scored slightly higher on our concise Machiavellianism scale than did women, $t(242) = 1.56, ns, d = 0.21$. The sex difference in Machiavellianism has proven to be somewhat elusive (e.g., Jonason, Li, & Buss, 2010). Consistent with Study 1, men scored higher on these measures than did women, on average.

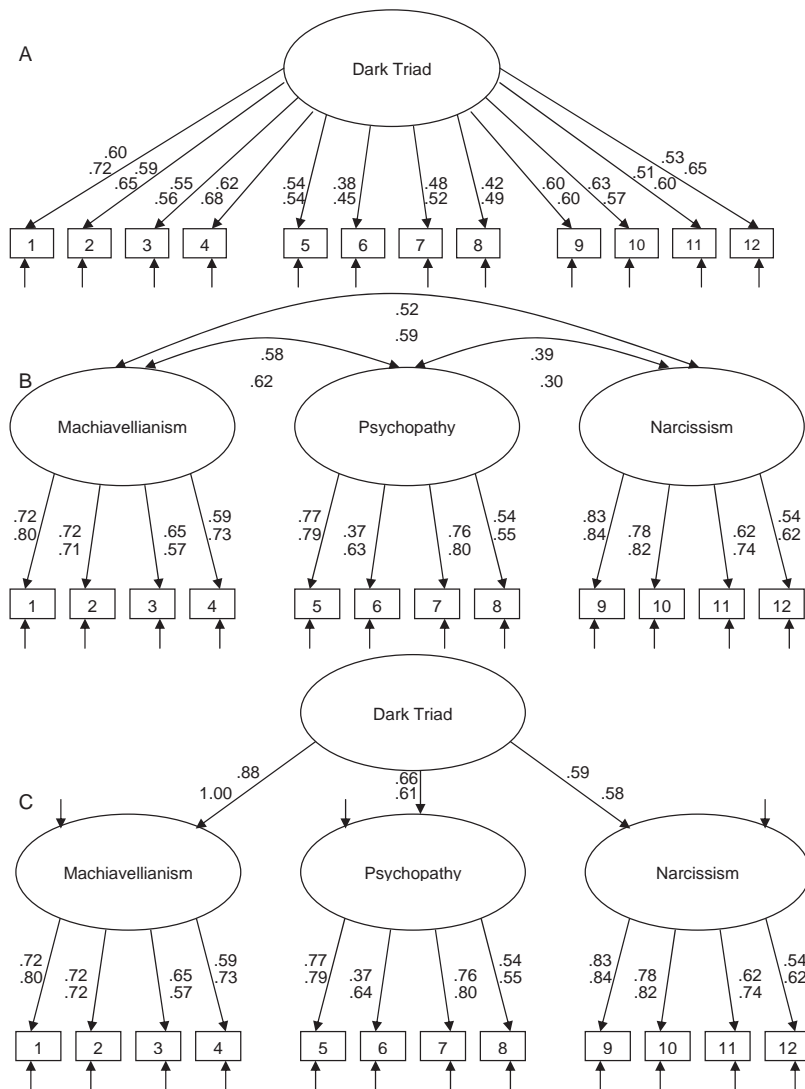


Figure 1. One-dimensional (A), three-dimensional (B), and hierarchical (C) confirmatory factor analyses of the Dark Triad Dirty Dozen items in Studies 2 (higher numbers, $N = 246$) and 4 (lower numbers, $N = 470$). All loadings and correlations were significant ($ps < .05$).

Study 3

We conducted Study 3 with three purposes in mind. First, we sought to assess the test–retest reliability of the Dark Triad Dirty Dozen and its three component scales. Second, we sought to assess the temporal reliability of each of the Dirty Dozen’s 12 items. Third, we sought to expand our understanding of the Dirty Dozen’s convergent and discriminant validity by correlating it with measures of aggression and self-esteem. Because aggression is positively related to psychopathy (Jones & Paulhus, 2010) and narcissism (Bushman & Baumeister, 1998; Bushman et al., 2009; Donnellan, Trzesniewski, Robins, Moffitt, & Caspi, 2005; Twenge & Campbell, 2003; Webster, 2006), and because Machiavellians may occasionally use aggression to manipulate others, we predicted that the Dark Triad Dirty Dozen would be positively correlated with self-reported aggression. Such predictions are also consistent with the original formulation of the Dark Triad in that they share an aggressive core (Paulhus & Williams, 2002).

We remained agnostic, however, as to what specific components of the Dark Triad would correlate most strongly with different measures of aggression (i.e., anger, hostility, and verbal and physical aggression). In part, this is because it appears that those who score high on the Dark Triad use tactics of aggression and manipulation as expressions of a “whatever works” strategy for dealing with conspecifics (Jonason, 2010). Therefore, we make no specific predictions about which types of aggressiveness will be correlated with the Dark Triad.

In contrast, showing evidence of discriminant validity, we predicted that the Dark Triad Dirty Dozen would be mostly uncorrelated with measures of self-esteem level and stability, aside from research suggesting that narcissism might represent unstable inflated self-views (Bushman & Baumeister, 1998; Rhodewalt, Madrian, & Cheney, 1998; but see also Webster, Kirkpatrick, Nezlek, Smith, & Paddock, 2007). Such a prediction would replicate results from Study 2 but extends those results by considering instability and change over time, a more psychometrically robust measurement technique for self-esteem.

Method

Participants and procedure. Participants were a convenience sample of 96 undergraduates enrolled in social or evolutionary psychology classes or a research laboratory group in the Southeastern United States. Each participant was asked to complete measures in class or lab once a week for three weeks. In all, 81 (84%), 79 (82%), and 76 (79%) participants completed questionnaires during Weeks 1, 2, and 3, respectively. The modal and median participant completed questionnaires all 3 weeks ($M = 2.46$, $SD = 0.77$). In all, 16 (17%), 20 (21%), and 60 (62%) participants completed only one, only two, or all three sessions, respectively. Of the 96 participants, 60 were women and 36 were men; the modal and median age was 20 years ($M = 20.44$, $SD = 1.43$), ranging from 18 to 25. Among the 60 participants who provided data for all 3 weeks, 37 were women and 23 were men; the modal and median age was 20 years ($M = 20.25$, $SD = 1.20$), ranging from 18 to 23.

Measures. All measures used a response scale from 1 (*strongly disagree*) to 9 (*strongly agree*). Study 3 used the same 12-item Dark Triad Dirty Dozen developed previously in Studies 1 and 2.

We measured self-esteem level with the Single-Item Self-Esteem Scale (“I have high self-esteem”; Robins, Hendlin, & Trzesniewski, 2001). Self-esteem instability was measured with two scales: (a) the five-item Labile Self-Esteem Scale (Dykman, 1998; $\alpha_s = .88, .89, .90$) and (b) the five-item Stability of Self Scale (Rosenberg, 1965; reverse-scored to reflect self-esteem instability; $\alpha_s = .86, .86, .84$). The Labile Self-Esteem Scale contained such items as “I notice that how good I feel about myself changes from day to day (or hour to hour)” and “I’m often feeling good about myself one minute, and down on myself the next minute,” whereas the Stability of Self Scale contained such items as “My opinion of myself tends to change a good deal.” Both self-esteem instability scales contained a reverse-scored item.

Aggression was measured using a brief, 12-item version of the Aggression Questionnaire (“I am an even-tempered person” [reverse-scored]; A. H. Buss & Perry, 1992; $\alpha_s = .75, .83, .82$) that has been used effectively in prior research (e.g., Webster, 2006, 2007; Webster et al., 2007). The brief Aggression Questionnaire uses the three highest loading items from each of four subscales found in A. H. Buss and Perry’s (1992) original article: Physical Aggression ($\alpha_s = .85, .85, .89$), Verbal Aggression ($\alpha_s = .55, .62, .68$), Hostility ($\alpha_s = .39, .48, .49$), and Anger ($\alpha_s = .65, .81, .84$).

Results and Discussion

Test–retest reliability. First, we examined the test–retest reliability of the Dark Triad Dirty Dozen and its three components in the traditional way by correlating Time 1 with Time 2 scores (r_{12}), Time 2 with Time 3 scores (r_{23}), and Time 1 with Time 3 scores (r_{13}); these correlations are bolded in Table 5. The average test–retest correlation (i.e., the Fisher’s r -to- z -based mean of r_{12} , r_{23} , and r_{13}) was .89 for the Dirty Dozen and ranged from .76 to .87 for its three subscales (see Table 6, first column).

Second, we examined the corrected test–retest reliabilities using Heise’s (1969) formula for three time points, $r_{xx} = (r_{12}r_{23})/r_{13}$, which estimates test–retest reliability independent of change-over-time effects. The corrected test–retest correlation was .91 for the Dark Triad Dirty Dozen and ranged from .71 to .88 for its subscales (see Table 6, second column).

Third, we used a series of structural equation measurement models to examine the test–retest reliabilities among latent variables for each measure at each time point for the 60 participants with complete data. For example, the measurement model for Machiavellianism was structurally identical to the model shown in Figure 1B, but with Machiavellianism at Times 1, 2, and 3 being the three latent variables. In contrast, the measurement model for the Dark Triad Dirty Dozen was a second-order model in which a latent variable for the Dark Triad at each time point was specified by the three latent factors of its subscales at that time point. We then took the resulting correlations among the three latent factors and submitted them to Heise’s (1969) formula. The corrected test–retest correlations based on latent variables was .94 for the Dark Triad Dirty Dozen and ranged from .67 to .92 for its subscales (see Table 6, third column).

Fourth, we reran the analyses above using data from all participants ($N = 96$) instead of those with complete data ($n = 60$), because maximum likelihood procedures provide sound estimates for data missing at random. The corrected test–retest correlations based on latent variables using all participants was .97 for the Dark

Table 5
Temporal Correlation Matrix for the Dark Triad Dirty Dozen and Its Three Subscales

Variable by Week	Dark Triad			Machiavellianism			Psychopathy			Narcissism		
	W1	W2	W3	W1	W2	W3	W1	W2	W3	W1	W2	W3
Dark Triad												
Week 1	(.78)											
Week 2	.86	(.83)										
Week 3	.87	.92	(.85)									
Machiavellianism												
Week 1	.87	.76	.77	(.81)								
Week 2	.81	.89	.81	.86	(.82)							
Week 3	.80	.83	.91	.85	.87	(.87)						
Psychopathy												
Week 1	.50	.42	.39	.30	.42	.29	(.44)					
Week 2	.40	.60	.52	.37	.55	.48	.66	(.59)				
Week 3	.43	.54	.56	.29	.46	.42	.77	.83	(.64)			
Narcissism												
Week 1	.73	.62	.65	.42	.40	.46	.04	-.04	.04	(.81)		
Week 2	.62	.70	.66	.39	.40	.45	-.01	-.01	.09	.85	(.88)	
Week 3	.71	.69	.77	.51	.47	.57	-.01	.02	.07	.87	.89	(.86)
<i>M</i>	4.08	4.03	4.03	4.22	4.17	4.18	2.42	2.40	2.51	5.62	5.53	5.41
<i>SD</i>	1.07	1.14	1.21	1.84	1.79	1.93	1.01	1.08	1.12	1.53	1.69	1.62

Note. *N* = 60. *r*s > .25 are significant at *p* < .05. Test-retest reliability correlations are bolded. Internal consistency coefficients (α s) appear along the diagonal. The α s for Psychopathy improved to .60, .63, and, .74, respectively, when Item 6 was omitted.

Triad Dirty Dozen and ranged from .79 to .91 for its subscales (see Table 6, fourth column).

Fifth, using Fisher's *r*-to-*z* transformation, we averaged across these four types of test-retest correlations to produce grand mean test-retest correlations. The grand mean test-retest correlation was .93 for the Dark Triad Dirty Dozen and ranged from .74 to .89 for its subscales (see Table 6, fifth column).

Item-level temporal reliability. We examined item-level temporal reliability using Cronbach's alpha (α) for participants with complete data (*N* = 60). That is, we examined the reliability of each of the Dirty Dozen items across three time points to assess each item's stability over time. Items with higher α s show greater temporal reliability, whereas items with lower α s show less. Item-level temporal reliability α s are shown in the rightmost column of Table 2, and ranged from .72 (Item 6) to .94 (Item 3) with a

Fisher's *r*-to-*z*-based mean of .90. The mean item-level temporal reliability α s for the Machiavellianism, Psychopathy, and Narcissism subscales were .92, .84, and .92, respectively.

Convergent and discriminant validity. Data were averaged across weeks for validity analyses (*N* = 60). As expected, neither the Dark Triad Dirty Dozen nor its constituent subscales were significantly correlated with the Single-Item Self-Esteem Scale (see Table 4). The Labile Self-Esteem Scale did not correlate significantly with the Dark Triad Dirty Dozen or its subscales; however, the Stability of Self Scale—recoded to reflect instability—was positively correlated with the Dark Triad, but this was driven largely by its positive correlation with Psychopathy. Overall, the Dark Triad showed no pattern of consistent correlation with either self-esteem level or self-esteem stability. In other words, these measures of self-esteem

Table 6
Test-Retest Reliability Correlations for the Dark Triad Dirty Dozen and Its Three Subscales Across Three Time Points (3 Weeks) Using Four Different Methods

Variable	Mean of 3 Uncorrected <i>N</i> = 60	Corrected Measured <i>N</i> = 60	Latent		Grand <i>M</i>
			<i>N</i> = 60	<i>N</i> = 96	
Machiavellianism	.86	.88	.92	.91	.89
Psychopathy	.76	.71	.67	.79	.74
Narcissism	.87	.87	.89	.90	.88

Note. Mean of 3 = Fisher's *r*-to-*z*-based mean of r_{12} , r_{23} , and r_{13} ; Corrected = corrected test-retest reliabilities using Heise's (1969) formula for three time points; $r_{xx} = (r_{12}r_{23})/r_{13}$, which estimates test-retest reliability independent of change-over-time effects; Latent = correlations among latent variables based on structural equation measurement models with maximum likelihood estimation; Grand *M* = Fisher's *r*-to-*z* based mean across the four types of test-retest correlations to produce grand mean test-retest correlations.

showed some discriminant validity with the Dark Triad Dirty Dozen.

In contrast, the Dark Triad Dirty Dozen showed some convergent validity with measures of aggression (see Table 4). As expected, the 12-item Dark Triad and the 12-item Aggression Questionnaire showed a strong correlation of .51, suggesting that people who score high on the Dark Triad may also use aggression to get what they want. Among subscales, Machiavellianism was positively related to physical aggression, verbal aggression, and hostility, but not anger—a pattern shared with the Dark Triad Dirty Dozen overall. Psychopathy, in contrast, was positively related only to physical and verbal aggression. Narcissism, on the other hand, was positively related only to hostility. Overall, the 12-item Aggression Questionnaire was positively related to Machiavellianism and Psychopathy, but not Narcissism. Nevertheless, the abbreviated Dark Triad and Aggression Questionnaires showed a fairly consistent pattern of convergent validity.

Because men scored higher on the Aggression Questionnaire than did women, $t(58) = 2.76, p < .01, d = 0.73$, especially on the Physical Aggression subscale, $t(58) = 5.26, p < .01, d = 1.38$, and because prior research has shown reliable sex differences in self-reported (Webster, 2006; Webster et al., 2007) and behavioral (Eagly & Steffen, 1986) aggression, we also ran the above correlations controlling for participant sex (see Table 4). Controlling for sex had no effect on the pattern of significant correlations with two exceptions: the correlation between Machiavellianism and Physical Aggression became nonsignificant, and the correlation between the Dark Triad Dirty Dozen and Anger became significant.

Sex differences. Tests for sex differences revealed that men scored higher than women on our concise measures of the Dark Triad, $t(58) = 2.63, p = .01, d = 0.69$, and Machiavellianism, $t(58) = 3.00, p < .01, d = 0.79$; men scored slightly higher than women on psychopathy, $t(58) = 1.34, ns, d = 0.35$, and narcissism, $t(58) = 1.29, ns, d = 0.34$. Consistent with Studies 1 and 2, men scored higher on these measures than did women, on average.

Study 4

A key limitation of the Dark Triad Dirty Dozen is the convoluted, double-barreled phrasing of Item 6, “I tend to not be too concerned with morality or the morality of my actions.” Double-barreled items should generally be avoided because respondents could hold different views about different topics within a single item (Simms & Watson, 2007). For example, it is possible for one to be concerned with morality while not being concerned with the morality of one’s actions, and such a person would have difficulty responding to Item 6. We suspect that this problem resulted in lower internal consistency coefficients (α s) for the psychopathy subscale compared with the narcissism and Machiavellianism subscales. To address this concern, we rephrased Item 6 to read, “I tend to be unconcerned with the morality of my actions.” The purpose of Study 4 was to examine the psychometric properties of the Dark Triad Dirty Dozen using this new, simplified Item 6. We expected that this new item would improve the psychopathy subscale’s internal consistency while replicating the Dark Triad Dirty Dozen’s three-dimensional factor structure.

Method

Participants and procedure. Participants were 470 undergraduate psychology students (157 men, 312 women, 1 unknown gender) aged 17 years ($n = 2$) to “26 and up” ($n = 3$) years (mode = 18, $Mdn = 19, M = 19.00, SD = 1.30$) from the Southeastern United States who received course credit for completing an online prescreening survey consisting of multiple questionnaires.

Measures. Study 4 used the same 12-item Dark Triad Dirty Dozen developed in Studies 1–3, with the exception that Item 6 was streamlined to read, “I tend to be unconcerned with the morality of my actions.” Participants responded using a scale from 1 (*disagree strongly*) to 9 (*agree strongly*).

Results and Discussion

Factor structure. Following recommendations by John and Soto (2007), we present item-level correlations and descriptive statistics in Table 7. This interitem correlation matrix shows the clustering of the Dark Triad Dirty Dozen items into their respective dimensions.

Consistent with Studies 1 and 2, we conducted a PCA (see Table 8) and a CFA (see Figure 1) on the Dark Triad Dirty Dozen. As predicted, the PCA produced a clear three-factor solution: Machiavellianism, Psychopathy, and Narcissism (see Table 8). The correlations among the three rotated factors and among the three subscales created from the unweighted items were moderately strong (see Table 9). Overall, the Dark Triad Dirty Dozen and its subscales achieved good internal consistency. Revising Item 6 in Study 4 resulted in a markedly stronger α for the psychopathy subscale (.77 vs. .63, .69, .44, .59, and .64 in Studies 1–3).

Replicating Studies 1 and 2, the single-factor measurement model (see Figure 1A) fit the data poorly, $\chi^2(54) = 822.52, p < .01, \chi^2/df = 15.23$; NFI = .64, CFI = .66; RMSEA = .17, 90% CI = .16–.18, $p_{\text{close fit}} < .01$. Because they are structurally equivalent, the three-dimensional (see Figure 1B) and hierarchical (see Figure 1C) models produced identical fits to the data, which were reasonable, $\chi^2(51) = 192.02, p < .01, \chi^2/df = 3.76$; NFI = .92, CFI = .94; RMSEA = .08, 90% CI = .06–.09, $p_{\text{close fit}} < .01$. The three-dimensional and hierarchical models fit the data significantly better than the one-dimensional model, $\Delta\chi^2(3) = 630.50, p < .01$. Nevertheless, the hierarchical model produced one negative variance estimate (the disturbance for the latent Machiavellianism factor). When this variance was constrained to equal zero, the fit remained good, $\chi^2(52) = 194.21, p < .01, \chi^2/df = 3.74$; NFI = .92, CFI = .94; RMSEA = .08, 90% CI = .06–.09, $p_{\text{close fit}} < .01$, and this constraint did not significantly degrade the model’s fit, $\Delta\chi^2(1) = 2.19, p = .14$.

Sex differences. Tests for sex differences revealed that men scored higher than women on our concise measures of the Dark Triad, $t(466) = 2.47, p = .01, d = 0.23$, and psychopathy, $t(466) = 4.99, p < .01, d = 0.46$; men scored slightly higher than women on Machiavellianism, $t(466) = 0.58, ns, d = 0.05$, and narcissism, $t(466) = 0.97, ns, d = 0.09$. These results are consistent with Studies 1–3 and prior work.

Table 7
Study 4 Dark Triad Dirty Dozen Item-Level Correlations and Descriptive Statistics

Item	Machiavellianism				Psychopathy				Narcissism			
	1	2	3	4	5	6	7	8	9	10	11	12
Machiavellianism												
Item 1	—											
Item 2	.58	—										
Item 3	.46	.45	—									
Item 4	.59	.49	.37	—								
Psychopathy												
Item 5	.33	.35	.19	.47	—							
Item 6	.25	.32	.18	.34	.55	—						
Item 7	.33	.32	.17	.44	.64	.49	—					
Item 8	.33	.31	.23	.34	.37	.27	.51	—				
Narcissism												
Item 9	.35	.32	.37	.28	.18	.13	.15	.24	—			
Item 10	.32	.29	.37	.29	.10	.12	.11	.19	.71	—		
Item 11	.38	.29	.33	.30	.20	.17	.20	.22	.63	.58	—	
Item 12	.48	.36	.37	.40	.27	.25	.25	.28	.45	.51	.48	—
<i>M</i>	3.80	3.91	4.81	2.57	2.27	1.92	2.38	3.29	5.64	5.09	5.45	3.31
<i>SD</i>	2.04	2.14	2.30	1.82	1.78	1.52	1.84	2.29	2.26	2.16	2.31	2.05

Note. Because of sporadic missing data, *N*s ranged from 466 to 469. All correlations were significant at $p < .05$. Intrasubscale correlations are bolded.

General Discussion

There is a growing trend in psychological assessment to create concise measures of core personality traits. Concise measures are appealing because they take less time to complete than more protracted personality inventories by eliminating item redundancy and thus reducing participant fatigue and frustration (Burisch, 1984, 1997; Saucier, 1994). Moreover, researchers have been demanding and developing more efficient measures of traditional scales for use in a variety of settings where a premium is placed on time or the number of items used (e.g., daily diary studies, experience sampling studies, mass-testing and prescreening sessions, field studies, research using special populations). In the present research, we developed and validated a concise measure of the

Dark Triad. These Dirty Dozen items had reasonable psychometric properties, showed acceptable convergent (e.g., NPI, Big Five, mating, aggression) and discriminant validity (e.g., self-esteem) with the other measures we examined and proved to be reliable over time and across a number of tests (e.g., corrected test-retest reliability). This single measure—the Dirty Dozen—provides a considerable improvement in efficiency compared to the 91-item, three-scale version of the Dark Triad (an 87% reduction in items). The Dirty Dozen version of the Dark Triad will not only reduce participant fatigue, it will also allow all three constructs to be measured using the same response scale format.

We found consistent evidence the Dirty Dozen measures the Dark Triad and shows correlation patterns with other personality

Table 8
Principal Components Analysis Using Oblique Rotation of and Item-Scale Correlations for the Dirty Dozen Dark Triad Items in Study 4

Item	Factor			
	M	P	N	<i>r</i>
1. I tend to manipulate others to get my way.	-.81	.04	.02	<i>.64</i>
2. I have used deceit or lied to get my way.	-.84	.04	-.08	<i>.57</i>
3. I have use flattery to get my way.	-.74	-.16	.12	<i>.51</i>
4. I tend to exploit others towards my own end.	-.64	.28	-.02	<i>.62</i>
5. I tend to lack remorse.	-.04	.83	-.02	<i>.50</i>
6. I tend to be unconcerned with the morality of my actions.	.04	.77	.00	<i>.42</i>
7. I tend to be callous or insensitive.	.02	.87	-.03	<i>.48</i>
8. I tend to be cynical.	-.07	.57	.13	<i>.46</i>
9. I tend to want others to admire me.	.02	-.01	.89	<i>.57</i>
10. I tend to want others to pay attention to me.	.01	-.08	.90	<i>.54</i>
11. I tend to seek prestige or status.	.03	.08	.83	<i>.57</i>
12. I tend to expect special favors from others.	-.22	.14	.55	<i>.61</i>

Note. Factor loadings $\geq |.30|$ are boldfaced. M = Machiavellianism; P = Psychopathy; N = Narcissism; *r* = corrected item-scale correlation (in italics).

Table 9
Study 4 Principal Component Correlation Matrix (Above the Diagonal), Zero-Order Correlations (Below the Diagonal), and Scale Reliability Coefficients (Along the Diagonal)

Variable	Dark Triad	Machiavellianism	Psychopathy	Narcissism
Dark Triad	(.86)			
Machiavellianism	.85**	(.79)	-.43	-.47
Psychopathy	.72**	.49**	(.77)	.24
Narcissism	.81**	.54**	.31**	(.84)
<i>M</i>	3.71	3.78	2.47	4.88
<i>SD</i>	1.30	1.63	1.44	1.80

Note. Alphas are presented in parentheses. Oblique rotation was used.
 ** $p < .01$.

traits in the nomological network. Specifically, the Dark Triad Dirty Dozen showed a consistent pattern of disagreeableness (Paulhus & Williams, 2002) and short-term mating (Jonason et al., 2009) across two studies and conscientiousness (Jonason, Li, & Teicher, in press), which may relate to a *fast* life strategy that underlies the nature of the Dark Triad (e.g., Jonason, Koenig, & Tost, 2010) in Study 2. In addition, after we corrected for measurement error, evidence suggests that the Dirty Dozen measures are negatively correlated with neuroticism and positively correlated with openness; both findings are consistent with work that suggests the Dark Triad reflects a latent dimension of social exploitation (Jonason et al., 2009), where this profile of lower order personality traits is expected (Jonason, Li, & Teicher, in press).

In line with this contention about social exploitation, we found that the Dark Triad tended to be correlated with measures of aggression in Study 3. Individuals may use tactics such as aggression to get what they want in life. One of the defining features of the Dark Triad is its link to aggressiveness (Paulhus & Williams, 2002). Indeed, new evidence suggests that one tactic used by those scoring high on the Dark Triad to get what they want is coercion (Jonason, 2010). However, we did not find much evidence for a narcissism–aggression correlation beyond narcissism’s positive correlation with hostility (cf. Bushman & Baumeister, 1998; Bushman et al., 2009; Donnellan et al., 2005; Jones & Paulhus, 2010; Twenge & Campbell, 2003; Webster, 2006). This may be because we did not include an ego threat (e.g., Bushman & Baumeister, 1998; Jones & Paulhus, 2010). Studies 2 (Self-Esteem Scale) and 3 (Single-Item Self-Esteem Scale) also show the lack of relationship between self-esteem and the Dark Triad. Self-esteem is not associated with the Dark Triad, despite the common conceptualization (e.g., Engler, 2009) that those who are, for instance, narcissistic may also have low self-esteem. Indeed, in Study 3, correlations between measures of self-esteem and the Dirty Dozen were some of the weakest we report. However, such evidence is inconsistent with research suggesting a possible link between the instability of self-esteem and narcissism (Bushman & Baumeister, 1998; Rhodewalt et al., 1998). It may be that the reduction of items from the 40-item NPI to the four items we used to measure narcissism caused one or more specific aspects of narcissism that relate to self-esteem instability to be lost (e.g., entitlement, grandiosity, superiority). This is a cost that researchers need to consider because the reduction of items in scales reduces some of narcissism’s heterogeneity.

Across numerous studies, men consistently scored higher than women did on the traditional measures of the Dark Triad (e.g., Jonason et al., 2009) and related measures like sociosexuality (e.g., Simpson & Gangestad, 1991). From an adaptationist perspective (D. M. Buss, 2009), men should benefit more from social exploitation (D. M. Buss & Duntley, 2008) and therefore should have higher scores on personality traits that reflect social exploitation. As measures that tap into this latent, exploitive psychology, our Dirty Dozen measures showed some sex differences across all studies, confirming that men tend to be more socially exploitive than women are through personality traits like the Dark Triad. Exploiting others may come at a higher cost for women than for men, because women are more dependent on social networks than men are on average (e.g., Jonason, Webster, & Lindsey, 2008).

Although some of the fit indexes could have been stronger (especially for the one-dimensional model), fit indexes do not have strict, nonarbitrary cutoff criteria (Fan & Sivo, 2007). Moreover, latent measures with fewer items or manifest indicators are expected to have more error than ones with more items, and thus, we did not anticipate the fit indexes to be overwhelmingly strong (Kline, 2000). Nevertheless, the measurement models fit the data well for the (structurally equivalent) three-dimensional and hierarchical models. The fact that the three-dimensional and hierarchical models of the Dark Triad Dirty Dozen fit the data better than the one-dimensional model is not surprising, because overall measures were derived from separate but related subscales (e.g., Bryant & Smith, 2001; Webster & Bryan, 2007). However, a one-dimensional model appears to behave as we would expect, and thus, it may prove useful as Jonason et al. (2009) argued.

The three-component measures of the Dark Triad Dirty Dozen were composed of only four items each, and as such, they are likely to have relatively low levels of internal consistency (Nunnally, 1978; Schmitt, 1996). Because coefficient α is a function, in part, of the number of items in a scale, the α s we reported were reasonable (Carmines & Zeller, 1979). Indeed, the overall Dirty Dozen measure had a higher degree of internal consistency than did the three subscales, consistent with this reasoning.

Perhaps related to measurement error, correlations between the original Dark Triad measures and our measures could have been stronger. For instance, the generation of new items, instead of cherry-picking good items from preexisting scales, may have failed to tap all the aspects of heterogeneous constructs like the Dark Triad. However, there is no agreement about the specific factor structures of measures of the Dark Triad. For instance, the

NPI has been treated as both a seven-dimensional (Raskin & Terry, 1988) and a four-dimensional (Emmons, 1987) construct. Future work should attempt to replicate our findings, in hopes of providing stronger evidence for convergent validity, perhaps examining the Dirty Dozen in relation to different factor structures of narcissism to see what aspects of each measure of the Dark Triad that it taps best. Overall, however, we feel that our measure provides a satisfactory compromise between precision and efficiency that are often at odds in measurement (e.g., Gorsuch & McFarland, 1989).

We believe that our concise Dark Triad measure—the Dirty Dozen—will have immediate applications in a variety of settings that value efficient measurement such as large-scale national or international surveys, prescreening packets, longitudinal studies, daily diary studies, experience-sampling studies, and anywhere else researchers may have limited time or funding. Our findings suggest that our concise measure, despite being called the Dirty Dozen, cleanly measures the latent constructs of the Dark Triad.

References

- Auerbach, J. S. (1984). Validation of two scales for narcissistic personality disorder. *Journal of Personality Assessment*, 48, 649–653.
- Benet-Martinez, V., & John, O. P. (1998). Los Cinco Grandes across cultures and ethnic groups: Multitrait-multimethod matrix analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology*, 75, 729–750.
- Bryant, F. B., & Smith, B. D. (2001). Refining the architecture of aggression: A measurement model for the Buss-Perry Aggression Questionnaire. *Journal of Research in Personality*, 35, 138–167.
- Burisch, M. (1984). Approaches to personality inventory construction: A comparison of methods. *American Psychologist*, 39, 214–227.
- Burisch, M. (1997). Test length and validity revisited. *European Journal of Personality*, 18, 303–315.
- Bushman, B. J., & Baumeister, R. F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence? *Journal of Personality and Social Psychology*, 75, 219–229.
- Bushman, B. J., Baumeister, R. F., Thomaes, S., Ryu, E., Beeger, S., & West, S. G. (2009). Looking again, and harder, for a link between low self-esteem and aggression. *Journal of Personality*, 77, 427–446.
- Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, 63, 452–459.
- Buss, D. M. (2009). How can evolutionary psychology explain personality and individual differences? *Perspectives in Psychological Science*, 4, 359–366.
- Buss, D. M., & Duntley, J. D. (2008). Adaptations for exploitation. *Group Dynamics*, 12, 53–62.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment: Quantitative applications in the social sciences series* (Vol. 17). Newbury Park, CA: Sage.
- Christie, R., & Geis, F. L. (1970). *Studies in Machiavellianism*. New York, NY: Academic Press.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.
- Comrey, A. L. (1973). *A first course in factor analysis*. New York, NY: Academic Press.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., Moffitt, T. E., & Caspi, A. (2005). Low self-esteem is related to aggression, antisocial behavior, and delinquency. *Psychological Science*, 16, 328–335.
- Dykman, B. M. (1998). Integrating cognitive and motivational factors in depression: Initial tests of a goal-orientation approach. *Journal of Personality and Social Psychology*, 74, 139–158.
- Eagly, A. H., & Steffen, V. J. (1986). Gender and aggressive behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin*, 100, 309–330.
- Emmons, R. A. (1987). Narcissism theory and measurement. *Journal of Personality and Social Psychology*, 52, 11–17.
- Engler, B. (2009). *Theories of personality*. Belmont, CA: Wadsworth.
- Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research*, 42, 509–529.
- Gorsuch, R. L., & McFarland, S. G. (1989). Intrinsic/Extrinsic measurement: I/E and single-item scales. *Journal for the Scientific Study of Religion*, 28, 348–354.
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37, 504–528.
- Heise, D. R. (1969). Separating reliability and stability in test-retest correlation. *American Sociological Review*, 34, 93–101.
- Hunter, J. E., Gerbing, D. W., & Boster, F. J. (1982). Machiavellianism beliefs and personality: Construct invalidity of the Machiavellianism dimensions. *Journal of Personality and Social Psychology*, 43, 1293–1305.
- Jackson, J. J., & Kirkpatrick, L. A. (2007). The structure and measurement of human mating strategies: Toward a multidimensional model of sociosexuality. *Evolution and Human Behavior*, 28, 382–391.
- Jakobwitz, S., & Egan, V. (2006). The dark triad and normal personality traits. *Personality and Individual Differences*, 2, 331–339.
- John, O. P., & Soto, C. J. (2007). The importance of being valid: Reliability and the process of construct validation. In R. W. Robins, R. C. Fraley, & R. F. Kruger (Eds.), *Handbook of research methods in personality psychology* (pp. 461–494). New York, NY: Guilford Press.
- Jonason, P. K. (2010). *Getting what one wants in life: The Dark Triad and tactics of influence*. Manuscript in preparation.
- Jonason, P. K., Koenig, B., & Tost, J. (2010). *Psychopathy links the Dark Triad to Life History Theory*. Manuscript submitted for publication.
- Jonason, P. K., Li, N. P., & Buss, D. M. (2010). The costs and benefits of the Dark Triad: Implications for mate poaching and mate retention tactics. *Personality and Individual Differences*, 48, 373–378.
- Jonason, P. K., Li, N. P., Webster, G. W., & Schmitt, D. P. (2009). The Dark Triad: Facilitating short-term mating in men. *European Journal of Personality*, 23, 5–18.
- Jonason, P. K., Li, N. P., & Teicher, E. A. (in press). Who is James Bond? The Dark Triad as an agentic social style. *Individual Differences Research*.
- Jonason, P. K., Webster, G. D., & Lindsey, A. E. (2008). Solutions to the problem of diminished social interaction. *Evolutionary Psychology*, 6, 637–651.
- Jones, D. N., & Paulhus, D. L. (2009). In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 93–108). New York, NY: Guilford Press.
- Jones, D. N., & Paulhus, D. L. (2010). Different provocations trigger aggression in narcissists and psychopaths. *Social Psychological and Personality Science*, 1, 12–18.
- Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). London, England: Routledge.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw Hill.
- Paulhus, D. L., Hemphill, J., & Hare, R. (in press). *Manual for the Self-Report Psychopathy Scale (SRP-III)*. Toronto, Canada: Multi-Health Systems.

- Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality, 36*, 556–563.
- Raskin, R. N., & Terry, H. (1988). A principal components analysis of the Narcissistic Personality Inventory and further evidence of its construct validity. *Journal of Personality and Social Psychology, 54*, 890–902.
- Rhodewalt, F., Madrian, J. C., & Cheney, S. (1998). Narcissism, self-knowledge organization, and emotional reactivity: The effect of daily experiences on self-esteem and affect. *Personality and Social Psychology Bulletin, 24*, 75–87.
- Robins, R. W., Hendlin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg self-esteem scale. *Personality and Social Psychology Bulletin, 27*, 151–161.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Saucier, G. D. (1994). Mini-markers: A brief version of Goldberg's unipolar Big-Five Markers. *Journal of Personality Assessment, 63*, 506–516.
- Schmitt, N. (1996). Uses and abuses of coefficient alphas. *Psychological Assessment, 8*, 350–353.
- Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences, 25*, 167–177.
- Seemann, E. A., Buboltz, W. C., Jr., Thomas, A., Soper, B., & Wilkinson, L. (2005). Normal personality variables and their relationship to psychological reactance. *Individual Differences Research, 3*, 88–98.
- Simms, L. J., & Watson, D. (2007). The construct validation approach to personality scale construction. In R. W. Robins, R. C. Fraley, & R. F. Kruger (Eds.), *Handbook of research methods in personality psychology* (pp. 240–258). New York, NY: Guilford Press.
- Simpson, J., & Gangestad, S. (1991). Individual differences in sociosexuality: Evidence for convergent and discriminant validity. *Journal of Personality and Social Psychology, 60*, 870–883.
- Twenge, J. M., & Campbell, W. K. (2003). "Isn't it fun to get the respect that we're going to deserve?" Narcissism, social rejection, and aggression. *Personality and Social Psychology Bulletin, 29*, 261–272.
- Webster, G. D. (2006). Low self-esteem is related to aggression, but especially when controlling for gender: A replication and extension of Donnellan et al. (2005). *Representative Research in Social Psychology, 29*, 12–18.
- Webster, G. D. (2007). Is the relationship between self-esteem and physical aggression necessarily U-shaped? *Journal of Research in Personality, 41*, 977–982.
- Webster, G. D., & Bryan, A. (2007). Sociosexual attitudes and behaviors: Why two factors are better than one. *Journal of Research in Personality, 41*, 917–922.
- Webster, G. D., Kirkpatrick, L. A., Nezlek, J. B., Smith, C. V., & Paddock, E. L. (2007). Different slopes for different folks: Self-esteem instability and gender as moderators of the relationship between self-esteem and attitudinal aggression. *Self and Identity, 6*, 74–94.
- Wilson, D. S., Near, D., & Miller, R. R. (1996). Machiavellianism: A synthesis of the evolutionary and psychological literatures. *Psychological Bulletin, 119*, 285–289.

Received September 8, 2009

Revision received February 8, 2010

Accepted February 10, 2010 ■