



## Short Communication

# Putting the “IRT” in “Dirty”: Item response theory analyses of the Dark Triad Dirty Dozen—An efficient measure of narcissism, psychopathy, and Machiavellianism

Gregory D. Webster<sup>a</sup>, Peter K. Jonason<sup>b,\*</sup>

<sup>a</sup>Department of Psychology, University of Florida, P.O. Box 112250, Gainesville, FL 32611-2250, United States

<sup>b</sup>School of Social Sciences and Psychology, University of Western Sydney, Milperra, NSW 2214, Australia

## ARTICLE INFO

## Article history:

Received 31 May 2012

Received in revised form 10 August 2012

Accepted 22 August 2012

Available online 16 September 2012

## Keywords:

Dark Triad

Narcissism

Psychopathy

Machiavellianism

Item response theory

## ABSTRACT

The Dark Triad Dirty Dozen is a new, concise personality inventory designed to measure individual differences in narcissism, psychopathy, and Machiavellianism in sub-clinical populations. To date, the psychometric work on this measure has used classical test theory to assess its validity, reliability, and structure. In two independent samples ( $N = 1014$ ), we provide the first analyses of this measure using item response theory. In both samples, the Dark Triad Dirty Dozen efficiently recovered information and showed some evidence of differential scale functioning by sex, with a lower endorsement threshold for men compared to women. Results shed new light on the measurement of these “dark” personality traits.

© 2012 Elsevier Ltd. All rights reserved.

## 1. Introduction

Research on the Dark Triad traits—narcissism, psychopathy, and Machiavellianism—is growing (Hodson, Hogg, & MacInnis, 2009; Jonason, Valentine, Li, & Harbeson, 2011; Jones & Paulhus, 2010). To facilitate research on these traits, Jonason and Webster (2010) developed a 12-item measure called The Dark Triad Dirty Dozen (DTDD). The DTDD provides researchers with a simple, efficient alternative to the traditional 91-item Dark Triad measures, which includes three separate measures with differing response formats and a complex scoring system. The DTDD has already been used in many contexts, providing insight into mate preferences (Jonason et al., 2011), self-control (Jonason & Tost, 2010), life history theory (Crysel, Crosier, & Webster, 2012; Jonason, Koenig, & Tost, 2010; Jonason & Webster, 2012), and social influence (Jonason, Slomski, & Partyka, 2012; Jonason & Webster, 2012). Jonason and colleagues have also shown that men tend to score higher than women on Dark Triad traits (see also Cale & Lilienfeld, 2002).

To show the DTDD's validity as a measure of its three purported constructs, Jonason and Webster (2010) used classical test theory (CTT) methods (e.g., factor analyses). Showing convergent validity, the DTDD traits positively correlated with their respectively longer measures. Factor analyses of the DTDD revealed a

three-dimensional structure that can form a global, second-order factor (Jonason, Kavanagh, Webster, & Fitzgerald, 2011; Jonason, Li, Webster, & Schmitt, 2009). A key limitation of prior research, however, is it has relied solely on CTT assessment techniques. In contrast, item response theory (IRT) offers some key improvements over traditional CTT; IRT is more analytically flexible and gives researchers more information than CTT (Morizot, Ainsworth, & Reise, 2007). The present study is the first to go beyond CTT by using IRT to assess the DTDD.

### 1.1. Item response theory

IRT is a family of psychometric methods for empirically evaluating items and the latent traits they measure (Morizot, Ainsworth, & Reise, 2007; Simms & Watson, 2007). IRT has been adopted in multiple empirical research settings such as educational assessment-based computerized adaptive testing. It has enjoyed increasing popularity in scale construction and optimization in psychology, including self-report measures in social-personality psychology (Ackerman, Donnellan, & Robins, 2012; Fraley, Waller, & Brennan, 2000; Webster & Crysel, 2012).

In the present study, we used IRT models related to the two-parameter logistic model (2PLM), which uses parameters for item difficulty ( $\beta$ ) and item discrimination ( $\alpha$ ). Item *difficulty* reflects the amount of the latent trait necessary to have a 50% chance of endorsing the item. Item *discrimination* represents the degree to

\* Corresponding author.

E-mail address: [p.jonason@uws.edu.au](mailto:p.jonason@uws.edu.au) (P.K. Jonason).

**Table 1**  
The Dark Triad Dirty Dozen Items and Their Respective Subscales.

Factor or item
<i>Machiavellianism</i>
1. I tend to manipulate others to get my way
2. I have used deceit or lied to get my way
3. I have used flattery to get my way
4. I tend to exploit others towards my own end
<i>Psychopathy</i>
5. I tend to lack remorse
6. I tend to be unconcerned with the morality of my actions
7. I tend to be callous or insensitive
8. I tend to be cynical
<i>Narcissism</i>
9. I tend to want others to admire me
10. I tend to want others to pay attention to me
11. I tend to seek prestige or status
12. I tend to expect special favors from others

which an item can differentiate between people with similar levels of the same latent trait. Another key concept in IRT is item- and scale-level *information*, which is related to the concept of precision-of-measurement reliability in CTT. We urge readers interested in IRT to consult relevant reviews (Morizot et al., 2007) and empirical examples (Ackerman et al., 2012; Fraley et al., 2000; Webster & Crysel, 2012).

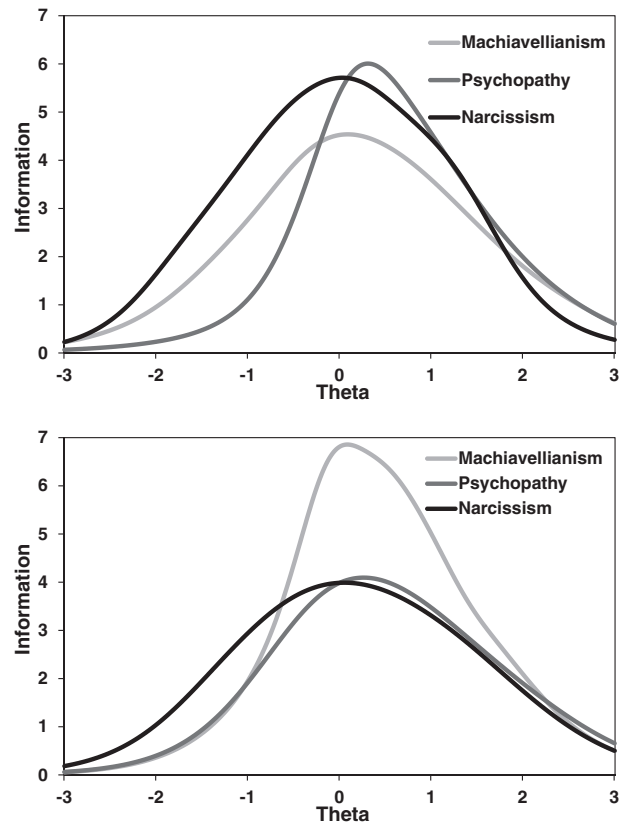
1.2. The present study

We had three goals. First, we determine the difficulty and discrimination parameter estimates for each item on its respective latent trait while accounting for covariance among the latent traits through the use of multidimensional IRT (MIRT; analogous to confirmatory factor analysis in CCT; see Finch, 2010; Muthén & Muthén, 2010). Second, we describe subscale-based information curves and the amount of total and per-item information that could be recovered for each trait measure. Third, we examine sex-based differential scale functioning (DSF).

2. Method

2.1. Sample 1

Participants were 470 undergraduates (66% women) enrolled in introductory psychology courses at a public university in the



**Fig. 1.** Scale Information Curves (SICs) for Dark Triad Dirty Dozen's three subscales (Sample 1, top; Sample 2, bottom).

southeastern US; ages ranged from 17 years to “26 and up” ( $M = 19.00, SD = 1.30$ ). Participants received course credit for completing an online prescreening survey that contained multiple questionnaires. Sample 1 represented secondary data analyses on data originally used in CTT-based factor analyses (Jonason & Webster, 2010).

2.2. Sample 2

Participants were 544 undergraduates (69% women) from a public university in the southeastern US; ages ranged from 17 to

**Table 2**  
Multidimensional item response theory parameter estimates for the 12 Dark Triad Dirty Dozen items by sample and subscales.

	Sample 1 (N = 470; 9-point scale)									Sample 2 (N = 544; 7-point scale)						
	$\alpha$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\beta_7$	$\beta_8$	$\alpha$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$
<i>Machiavellianism</i>																
1	1.41	-1.24	-0.59	0.06	0.40	0.88	1.56	2.35	2.64	2.34	-0.10	0.46	0.78	1.12	1.70	2.10
2	1.15	-1.35	-0.57	-0.10	0.34	0.85	1.47	2.37	2.72	1.11	-0.67	0.08	0.44	0.92	1.86	2.50
3	0.81	-1.99	-1.19	-0.75	-0.28	0.31	1.02	2.00	2.54	1.11	-0.70	-0.04	0.33	0.86	1.68	2.51
4	1.59	-0.27	0.30	0.77	1.10	1.53	2.03	2.67	2.79	1.55	-0.17	0.47	0.81	1.46	2.16	2.75
<i>Psychopathy</i>																
5	1.87	0.09	0.51	0.97	1.24	1.60	2.04	2.52	2.87	1.13	-0.29	0.42	0.85	1.63	2.23	2.59
6	1.39	0.29	0.91	1.53	1.88	2.17	2.39	2.72	3.00	1.24	-0.01	0.71	1.11	1.77	2.24	2.56
7	1.56	0.00	0.49	0.92	1.25	1.62	2.13	2.65	2.82	1.21	-0.39	0.34	0.72	1.28	1.96	2.69
8	0.74	-0.83	0.00	0.48	0.82	1.54	2.01	2.74	3.37	1.07	-0.65	0.03	0.43	1.25	1.95	2.62
<i>Narcissism</i>																
9	1.78	-1.82	-1.41	-0.95	-0.61	-0.16	0.27	1.00	1.34	1.17	-1.48	-0.80	-0.44	0.15	0.95	1.64
10	1.78	-1.78	-1.23	-0.74	-0.34	0.16	0.67	1.31	1.63	1.23	-1.32	-0.57	-0.19	0.47	1.29	2.19
11	1.32	-2.01	-1.44	-0.82	-0.48	-0.04	0.44	1.07	1.48	1.15	-1.09	-0.48	-0.11	0.64	1.34	1.93
12	0.97	-0.95	-0.30	0.32	0.80	1.46	2.09	2.62	2.83	1.11	-0.68	0.19	0.64	1.38	2.07	2.79

Note:  $\alpha$  = Discrimination parameter.  $\beta_{1-8}$  = Difficulty parameters.

**Table 3**  
Descriptive and information statistics for the Dark Triad Dirty Dozen (DTDD) and its subscales by participant sex and sample.

Subscale	Descriptives		Reliability		Information		Sex difference	
	Mean	SD	$\alpha$	MIC	Total	Per Item	<i>t</i>	<i>d</i>
Sample 1 ( <i>N</i> = 470; 9-point scale)								
<i>All</i>								
DTDD	3.71	1.30	.86	.34	46.85	3.90	2.47*	0.23
Machiavellianism	3.78	1.63	.79	.49	14.72	3.68	0.58	0.05
Psychopathy	2.47	1.44	.77	.48	14.06	3.52	4.99**	0.46
Narcissism	4.88	1.80	.84	.56	18.07	4.52	0.97	0.09
<i>Men</i>								
DTDD	3.92	1.37	.87	.36				
Machiavellianism	3.85	1.61	.79	.48				
Psychopathy	2.92	1.72	.81	.53				
Narcissism	5.00	1.79	.82	.53				
<i>Women</i>								
DTDD	3.61	1.25	.86	.34				
Machiavellianism	3.76	1.64	.79	.50				
Psychopathy	2.23	1.21	.72	.42				
Narcissism	4.83	1.81	.85	.59				
Sample 2 ( <i>N</i> = 544; 7-point scale)								
<i>All</i>								
DTDD	2.92	1.07	.87	.36	43.04	3.59	5.27***	0.45
Machiavellianism	2.69	1.32	.80	.51	16.66	4.16	3.62***	0.31
Psychopathy	2.54	1.20	.73	.41	12.55	3.14	4.79***	0.41
Narcissism	3.51	1.37	.78	.47	13.84	3.46	4.60***	0.40
<i>Men</i>								
DTDD	3.27	1.03	.84	.31				
Machiavellianism	3.00	1.38	.80	.52				
Psychopathy	2.90	1.21	.70	.37				
Narcissism	3.91	1.33	.79	.49				
<i>Women</i>								
DTDD	2.76	1.06	.87	.36				
Machiavellianism	2.56	1.27	.79	.49				
Psychopathy	2.38	1.17	.74	.41				
Narcissism	3.34	1.35	.77	.46				

Note. MIC = Mean interitem correlation.

\*  $p < .05$ .

\*\*  $p < .01$ .

\*\*\*  $p < .001$ .

50 years ( $M = 20.25$ ,  $SD = 4.70$ ). Participants received course credit for completing an online survey that contained multiple questionnaires. Sample 2 represented secondary data analyses used in a validation study using the HEXACO (Jonason & McCain, 2012).

### 2.3. Measures

We measured the Dark Triad traits using the DTDD (Jonason & Webster, 2010; Table 1). Participants used a response scale from 1 (*disagree strongly*) to 9 (*agree strongly*) in Sample 1 or from 1 (*disagree strongly*) to 7 (*agree strongly*) in Sample 2.

## 3. Results

### 3.1. Descriptive statistics

Table 3 shows the descriptive statistics for the three DTDD subscales by sample and sex. Means and SDs suggested that men scored significantly higher than women did on psychopathy and the composite Dark Triad in Sample 1, and on all Dark Triad traits in Sample 2.

### 3.2. Multidimensional item response theory analyses

#### 3.2.1. Discrimination, difficulty, and information

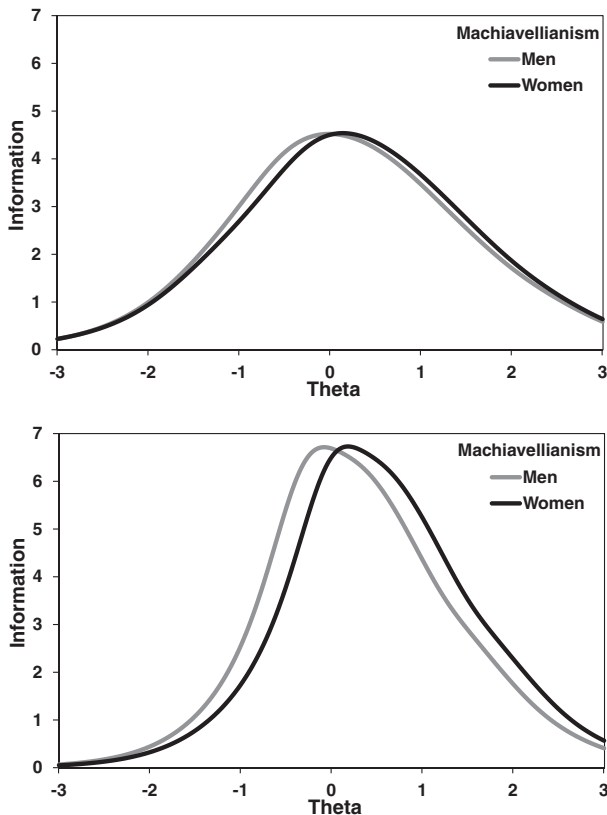
We ran a graded-response polytomous model (which is analogous to a 2PLM for dichotomous outcomes) on the 12 Dirty Dozen

items while allowing items to load on their respective latent subscales (dimensions). The discrimination ( $\alpha$ ) and difficulty ( $\beta$ s) parameters from these models appear in Table 2. Across samples, the 12 items had  $\alpha$ s ranging from 0.74 to 2.34, which suggested the items adequately discriminated among people along their respective latent traits. The eight (Sample 1) or six (Sample 2) difficulty parameters for each item ( $\beta$ s) suggested that DTDD items tended to be “difficult” to endorse on average, perhaps because each trait is considered socially undesirable. Narcissism items were easier to endorse, given the negative-to-positive crossover point in the difficulty parameters was near the midpoint of the latent variable; however, psychopathy and Machiavellianism items were more difficult to endorse across both samples.

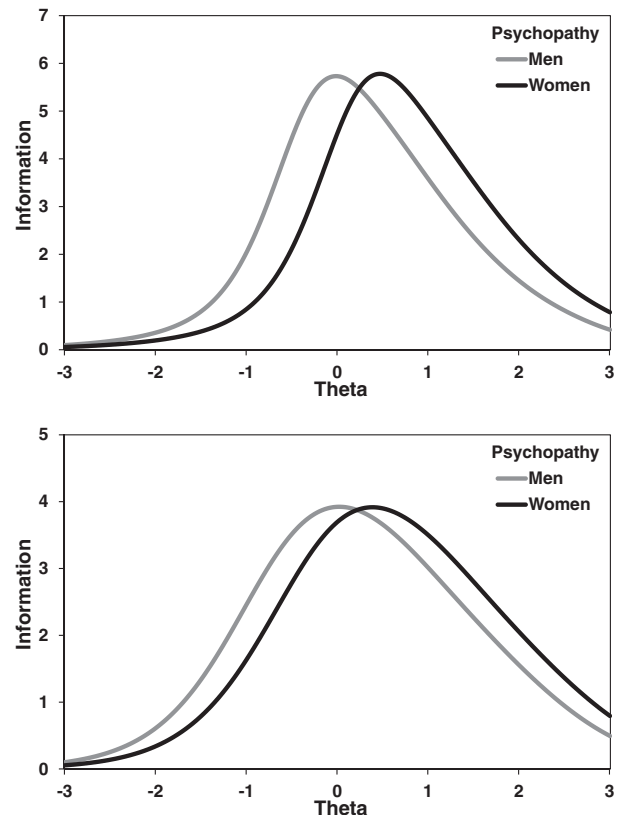
The corresponding scale information curves (SICs) for each subscale are overlaid in Fig. 1. Because SICs are an additive function of the number of items contributing to a given scale, they are impressive given that each subscale has only four items. Table 3 shows the total information (area under the curve—AUC) for each SIC for each subscale, and the total information per item. Recall that information in IRT is related to reliability in CTT. This relationship was confirmed in both samples: mean interitem correlations (MICs in Table 2) were correlated with SIC AUCs from the three subscales,  $rs(1) = .999$  and  $.947$  in Sample 1 and 2, respectively.

#### 3.2.2. Differential item and scale functioning

We next examined differential item functions (DIF) for Dirty Dozen subscales, and then constructed differential scale functions



**Fig. 2.** Constructed Differential Scale Functions (DSF) comparing men and women's scores on the Dark Triad Dirty Dozen's Machiavellianism subscale (Sample 1, top; Sample 2, bottom).



**Fig. 3.** Constructed Differential Scale Functions (DSF) comparing men and women's scores on the Dark Triad Dirty Dozen's psychopathy subscale (Sample 1, top; Sample 2, bottom).

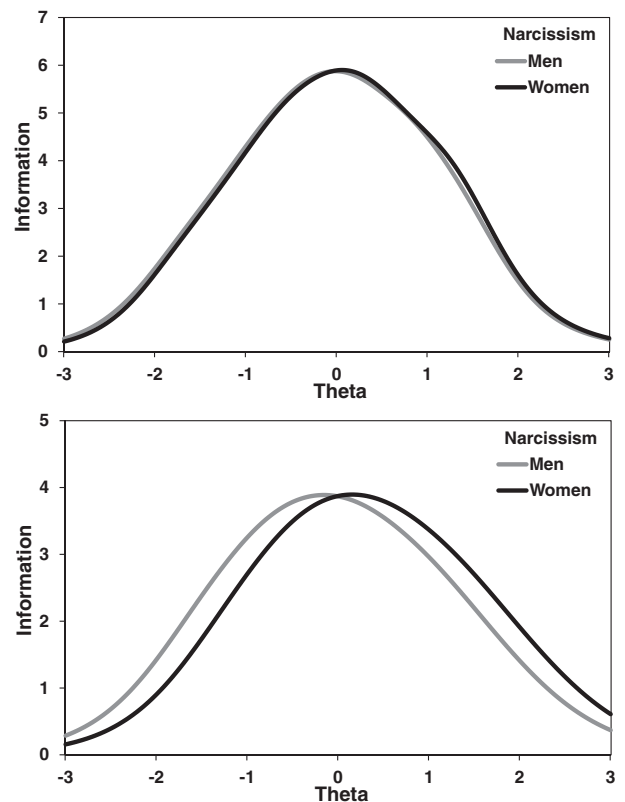
(DSF; see Morizot et al., 2007) for each subscale (Figs. 2–4). Figures. 2–4 show some interesting patterns of differential scale responding by sex (and sample)—key differences (and similarities) that can be masked using CTT alone.

Overall, the DSFs for all three traits (Figs. 2–4) showed substantial overlap between men and women in both samples, particularly for narcissism in Sample 1. Despite these similarities, men consistently had at least slightly lower thresholds than women for endorsing items of all three Dark Triad traits. For example, the largest gender difference was for psychopathy in Sample 1; men had a lower threshold for endorsing these items than women by about 0.5 SDs on a standardized latent measure ( $-0.01$  vs.  $0.47$ ). Thus, it was comparatively easier for men than women to endorse the Machiavellian items as a set. Nevertheless, we caution readers not to overlook the substantial overlap between the sexes, despite the small-to-moderate—but fairly consistent—sex differences.

#### 4. Discussion

We showed that all three DTDD subscales adequately and efficiently recovered information. Differential scale functioning analyses showed that men and women responded in largely similar ways to the DTDD subscales; however, men had slightly lower endorsement thresholds than women, on average, for each subscale. This provided a unique test to support prior results (Jonason & Webster, 2010; Jonason et al., 2009) that men score higher on these measures than women do.

Despite its strengths, the present research also has weaknesses. First, because the DTDD focuses on sub-clinical levels of the Dark Triad traits, we sampled normative student populations; future work might benefit from assessing these traits in clinical or



**Fig. 4.** Constructed Differential Scale Functions (DSF) comparing men and women's scores on the Dark Triad Dirty Dozen's narcissism subscale (Sample 1, top; Sample 2, bottom).

incarcerated populations. Second, Samples 1 and 2 used different response scales (9- vs. 7-point) to measure the DTDD; however, research suggests that response scale ranges have little effect on measurement properties (Dawes, 2008).

This study is the first to provide IRT analyses of the DTDD. Overall, the DTDD showed good discrimination and difficulty parameters, and adequate information recovery. In addition, although men had slightly lower item endorsement thresholds than women, the overall sex difference was fairly small—the overlap in DSFs by sex was substantial. The present findings add to a growing literature on the Dark Triad and its constituent traits, and highlight the psychometric robustness of a new efficient measure—the Dark Triad Dirty Dozen.

### Authors' note

Different analyses of Sample 1 data appear in Jonason and Webster (2010) and Sample 2 data appear in Jonason and McCain (2012).

### References

- Ackerman, R. A., Donnellan, M. B., & Robins, R. W. (2012). An item response theory analysis of the narcissistic personality inventory. *Journal of Personality Assessment*, 94, 141–155.
- Cale, E. M., & Lilienfeld, S. O. (2002). Sex differences in psychopathy and antisocial personality disorder: A review and integration. *Clinical Psychology Review*, 22, 1179–1207.
- Crysel, L. C., Crosier, B. S., & Webster, G. D. (2012). The Dark Triad and risk behavior. *Personality and Individual Differences*, 54, 35–40. <http://dx.doi.org/10.1016/j.paid.2012.07.029>.
- Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point, and 10-point scales. *International Journal of Market Research*, 50, 61–77.
- Finch, H. (2010). Item parameter estimation for the MIRT model: Bias and precision of confirmatory factor analysis-based models. *Applied Psychological Measurement*, 34, 10–15.
- Fraley, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology*, 78, 350–365.
- Hodson, G., Hogg, S. M., & MacInnis, C. C. (2009). The role of “dark personalities” (narcissism, Machiavellianism, psychopathy), Big Five personality factors, and ideology in explaining prejudice. *Journal of Research in Personality*, 43, 686–690.
- Jonason, P. K., Kavanagh, P., Webster, G. D., & Fitzgerald, D. (2011). Comparing the measured and latent Dark Triad: Are three better than one? *Journal of Methods and Measurement in the Social Sciences*, 2, 28–44.
- Jonason, P. K., Koenig, B., & Tost, J. (2010). Living a fast life: The Dark Triad and Life History Theory. *Human Nature*, 21, 428–442.
- 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., & McCain, J. (2012). Using the HEXACO model to test the validity of the Dirty Dozen measure of the Dark Triad. *Personality and Individual Differences*, 53, 935–938. <http://dx.doi.org/10.1016/j.paid.2012.07.010>.
- Jonason, P. K., Slomski, S., & Partyka, J. (2012). The Dark Triad at work: How toxic employees get their way. *Personality and Individual Differences*, 52, 449–453.
- Jonason, P. K., & Tost, J. (2010). I just cannot control myself: The Dark Triad and self-control. *Personality and Individual Differences*, 49, 611–615.
- Jonason, P. K., Valentine, K. A., Li, N. P., & Harbeson, C. L. (2011). Mate-selection and the Dark Triad: Facilitating a short-term mating strategy and creating a volatile environment. *Personality and Individual Differences*, 51, 759–763.
- Jonason, P. K., & Webster, G. D. (2010). The Dirty Dozen: A concise measure of the Dark Triad. *Psychological Assessment*, 22, 420–432.
- Jonason, P. K., & Webster, G. D. (2012). A protean approach to social influence: Dark Triad Personalities and Social Influence Tactics. *Personality and Individual Differences*, 52, 521–526.
- Jones, D. N., & Paulhus, D. L. (2010). Different provocations trigger aggression in narcissists and psychopaths. *Social Psychological and Personality Science*, 1, 12–18.
- Morizot, J. M., Ainsworth, A. T., & Reise, S. P. (2007). Towards modern psychometrics: Application of item response theory models in personality research. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 407–423). Guilford Press.
- Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide, version 6*. Los Angeles: Muthén and Muthén.
- Simms, L. J., & Watson, D. (2007). The construct validation approach to personality scale construction. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 240–258). New York: Guilford.
- Webster, G. D., & Crysel, L. C. (2012). “Hit me, maybe, one more time”: Brief measure of impulsivity and sensation seeking and their prediction of blackjack bets and sexual promiscuity. *Journal of Research in Personality*, 46, 591–598. <http://dx.doi.org/10.1016/j.jrp.2012.07.001>.