



## UvA-DARE (Digital Academic Repository)

### Does response distortion statistically affect the relations between self-report psychopathy measures and external criteria?

Watts, A.L.; Lilienfeld, S.O.; Edens, J.F.; Douglas, K.S.; Skeem, J.L.; Verschuere, B.; LoPilato, A.C.

**DOI**

[10.1037/pas0000168](https://doi.org/10.1037/pas0000168)

**Publication date**

2016

**Document Version**

Final published version

**Published in**

Psychological Assessment

**License**

Article 25fa Dutch Copyright Act (<https://www.openaccess.nl/en/in-the-netherlands/you-share-we-take-care>)

[Link to publication](#)

**Citation for published version (APA):**

Watts, A. L., Lilienfeld, S. O., Edens, J. F., Douglas, K. S., Skeem, J. L., Verschuere, B., & LoPilato, A. C. (2016). Does response distortion statistically affect the relations between self-report psychopathy measures and external criteria? *Psychological Assessment*, 28(3), 294-306. <https://doi.org/10.1037/pas0000168>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

# Does Response Distortion Statistically Affect the Relations Between Self-Report Psychopathy Measures and External Criteria?

Ashley L. Watts and Scott O. Lilienfeld  
Emory University

John F. Edens  
Texas A&M University

Kevin S. Douglas  
Simon Fraser University

Jennifer L. Skeem  
University of California, Berkeley

Bruno Verschuere  
University of Amsterdam

Alexander C. LoPilato  
University of Georgia

Given that psychopathy is associated with narcissism, lack of insight, and pathological lying, the assumption that the validity of self-report psychopathy measures is compromised by response distortion has been widespread. We examined the statistical effects (moderation, suppression) of response distortion on the validity of self-report psychopathy measures in the statistical prediction of theoretically relevant external criteria (i.e., interview measures, laboratory tasks) in a large sample of offenders ( $N = 1,661$ ). We conducted 378 moderation and 378 suppression analyses to examine the response distortion hypothesis. The substantial majority of analyses (97% moderation, 83% suppression) offered no support for this hypothesis. Nevertheless, suppression analyses revealed consistent evidence that controlling for response distortion slightly increased the relations between the fearless dominance and coldheartedness features of psychopathy and maladaptive outcomes. Our findings are largely inconsistent with the popular notion that the validity of self-report psychopathy measures is markedly diminished by response distortion. Further research is necessary to determine whether these findings generalize to other populations or contexts.

**Keywords:** psychopathy, personality assessment, self-report, response bias, social desirability

**Supplemental materials:** <http://dx.doi.org/10.1037/pas0000168.supp>

Psychopathic personality, or psychopathy, is a personality condition characterized by a constellation of affective, interpersonal, and behavioral traits including superficial charm, grandiosity, manipulativeness, lack of empathy, guiltlessness, poor impulse control, lack of insight, and deceit (Cleckley, 1941; McCord & McCord, 1964). In his classic clinical descriptions, Cleckley depicted psychopathic individuals as hybrid creatures whose maladaptive personality is often masked by apparent normalcy: “The typical

psychopath will seem particularly agreeable and make a distinctly positive impression when he is first encountered. [. . .] He looks like the *real thing*” (p. 339, emphasis added). Owing to psychopathy’s relations with (a) grandiose self-concept, (b) lack of insight, and (c) conning and deceptive behaviors, many scholars have questioned psychopathic individuals’ ability to report accurately on their personality traits and behaviors, rendering self-report psychopathy instruments controversial (Hare, 1985). For example,

---

This article was published Online First June 29, 2015.

Ashley L. Watts and Scott O. Lilienfeld, Department of Psychology, Emory University; John F. Edens, Department of Psychology, Texas A&M University; Kevin S. Douglas, Department of Psychology, Simon Fraser University; Jennifer L. Skeem, School of Social Welfare and Goldman School of Public Policy, University of California Berkeley; Bruno Verschuere, Department of Psychology, University of Amsterdam; Alexander C. LoPilato, Department of Psychology, University of Georgia.

This research was supported in part by National Institute of Mental Health Grant RO1 MH63783-01A1 (Personality Features in Social Deviancy). We acknowledge and appreciate the assistance and cooperation of the following agencies in collecting data for this research (however, none of the opinions or conclusions expressed in this article reflects any official policy or position of any of these institutions): Drug Abuse Comprehensive

Coordinating Office (DACC), Tampa, Florida; Florida Department of Corrections; Gateway Foundation, Huntsville, Texas; Nevada Department of Prisons; Odyssey House, Salt Lake City, Utah; Operation PAR, Pinellas Park, Florida; Oregon Department of Corrections; Texas Department of Criminal Justice—Institutional Division; Utah Department of Corrections; Volunteers of America, Portland, Oregon; WestCare Nevada, Las Vegas, Nevada. Scott O. Lilienfeld is the coauthor of the Psychopathic Personality Inventory. John F. Edens receives royalties from the sale of the Personality Assessment Inventory computerized interpretive report used in prisons but does not benefit financially from the sale of the Personality Assessment Inventory itself.

Correspondence concerning this article should be addressed to Ashley L. Watts, Department of Psychology, Emory University, 36 Eagle Row, Atlanta, GA 30322. E-mail: [alwatts@emory.edu](mailto:alwatts@emory.edu)

psychopathic individuals' propensities toward grandiose self-concept and poor insight may result in sincere, but positively distorted, reports of their attributes (Lilienfeld & Fowler, 2006). Furthermore, their propensities toward dishonesty may lead them to consciously dissimulate on questionnaires. In this article, we systematically examine these widespread conjectures.

### The Assessment of Psychopathy

Factor analyses of the widely used Psychopathy Checklist-Revised (PCL-R; Hare, 1991/2003) and its derivatives suggest that psychopathy is underpinned by at least two broad higher order dimensions and up to four lower order dimensions (cf. Cooke & Michie, 2001, for an alternative three-factor structure). In this well-established two-factor model of psychopathy, Factor 1 comprises interpersonal and affective traits (e.g., superficial charm, lack of empathy, lack of remorse) and Factor 2 comprises antisocial behavior and lifestyle traits (e.g., irresponsibility, impulsivity). To date, much of the research on psychopathy has focused on prison and other forensic settings, and as a consequence has relied on the PCL-R, which was designed for these settings (Hare, 1991/2003). Nevertheless, given evidence of psychopathy's dimensionality (Edens, Marcus, Lilienfeld, & Poythress, 2006), there has been burgeoning interest in studying psychopathy in nonclinical (e.g., college, community) settings, which has spurred the development of several well-validated self-report psychopathy measures, including the Psychopathic Personality Inventory (PPI; Lilienfeld & Andrews, 1996) and the Levenson Self-Report Psychopathy Scale (LSRP; Levenson, Kiehl, & Fitzpatrick, 1995).

Somewhat similar to those on the PCL-R, factor analyses of the PPI and its revised version, the PPI-R, often reveal two higher order dimensions, fearless dominance (FD) and self-centered impulsivity (SCI). Fearless dominance assesses social and physical boldness and immunity to stress, whereas SCI measures a narcissistic tendency to exploit others, recklessness, and blame externalization (Benning, Patrick, Hicks, Blonigen, & Krueger, 2003). Like PCL-R Factor 1, FD assesses many of the interpersonal and affective features of psychopathy. In contrast, it places greater emphasis on social boldness, physical fearlessness, and other potentially adaptive traits. FD correlates only modestly with PCL-R Factor 1, suggesting that despite some overlap, these two dimensions detect a somewhat different set of traits relevant to psychopathy (Marcus, Fulton, & Edens, 2013; Miller & Lynam, 2012). Whereas FD is associated largely with psychologically healthy functioning (e.g., emotional stability, adjustment), SCI is associated with maladaptive functioning (e.g., externalizing behaviors, including antisocial behavior; Lilienfeld et al., 2012). Additionally, one of the eight PPI subscales, Coldheartedness (C), does not load highly on either FD or SCI and is sometimes used as a third indicator of psychopathy that assesses a lack of empathy, guiltlessness, and lovelessness. Like the PPI, the LSRP generally comprises two broad factors, Primary (Factor 1) and Secondary (Factor 2) psychopathy, which appear to be allied with largely maladaptive aspects of psychopathy (e.g., anger, impulsivity; McHoskey, Worzel, & Szyarto, 1998) and capture both Factor 1 and Factor 2 of the PCL-R (Sellbom, 2011).

### Response Distortion

Broadly construed, positive response distortion typically reflects the tendency to deny or minimize socially undesirable traits and to emphasize socially desirable ones, resulting in an overly positive or negative impression of oneself (Davis, Thake, & Weekes, 2012). Some authors have distinguished two variants of this form of underreporting of pathological features. One variant, sometimes termed "self-deception," ostensibly reflects a largely subtle form of positive response distortion in which individuals endorse unrealistic cognitive traits and behaviors. The other variant, sometimes termed "other deception" or "positive impression management," ostensibly reflects a largely overt form of positive response distortion in which individuals attempt to make themselves appear virtuous and moral (Davis et al., 2012; Paulhus & John, 1998). Some authors (e.g., Ray et al., 2012) have referred to the latter form of response distortion as "faking good," although the extent to which this response bias is invariably deliberate remains controversial. This latter response bias may occur within a parole evaluation, for instance, when an offender wants to increase the odds of release. Nevertheless, these two response biases tend to be at least moderately correlated (e.g., Sackeim & Gur, 1979), and most positive response distortion measures contain a mix of both biases (Davis et al., 2012). Moreover, in simulation studies, measures of both response biases have been found to be susceptible to instructions to fake good (de Vries, Zettler, & Hilbig, 2014). Although positive response distortion is clearly a concern in certain forensic contexts, it is hardly the only response bias that can compromise the validity of assessments. Instead, in other forensic contexts, response distortion can assume the form of overreporting "bad" features, loosely termed "faking bad" or malingering when presumed to be largely deliberate. This might occur in personal injury or criminal responsibility evaluations when appearing impaired is consistent with one's legal objectives.

We focus on response styles associated with underreporting and overreporting of socially desirable personality traits and allied behaviors as opposed to other potentially problematic response biases (e.g., inconsistent responding, acquiescence, extreme responding), as they present particular challenges to psychopathy assessment. We were particularly interested in overreporting, given psychopathic individuals' propensities toward grandiosity, lack of insight, and dishonesty in everyday life. Hence, there is reason to suspect that both forms of positive response distortion, namely self-deception and impression management, may be causes for concern in the self-report assessment of psychopathy. In addition, in light of psychopathy's relations with "faking bad" and malingering in some settings (e.g., forensic), we also examined indices of overreporting (cf. Niensten, Nentjes, Merckelbach, & Bernstein, 2015).

Response distortion on questionnaires is among the most widely studied topics in applied psychological assessment (McGrath, Mitchell, Kim, & Hough, 2010). Owing to this interest, a number of scales have been developed to detect socially desirable response styles as well as other forms of response distortion, several of which are embedded within frequently used measures of personality and psychopathology (i.e., the Personality Assessment Inventory [PAI], Morey, 1991; the Multidimensional Personality Questionnaire [MPQ], Tellegen, 1982; and the PPI).

The question of whether response biases adversely affect the validity of self-report personality or psychopathology measures in general has a long history. For example, McCrae and Costa (1992; cf. Ben-Porath & Waller, 1992) titled their article, “Social desirability scales: More substance than style,” and argued that extant research offered little or no support for the view that controlling statistically for social desirability measures increased the relations between self-reported personality traits and external criteria. To the contrary, they found that the validity of self-reported personality measures often *decreased* following statistical control for social desirability indices. Similarly, Ones, Viswesvaran, and Reiss (1996) found meta-analytically that a broad swath of social desirability measures, including indices of both ostensible self-deception and impression management, did not function as useful suppressors in the prediction of job performance. They found further that social desirability indices are imbued with genuine personality variance, namely low levels of neuroticism and conscientiousness (see also Li & Bagger, 2006; McCrae & Costa, 1983). For example, impression management indices may largely assess adequate adjustment and healthy self-control (Uziel, 2010).

More recently, a controversial but influential meta-analysis by McGrath et al. (2010) examined again the effects of response bias in applied assessment in the broader personality literature. Consistent with the Ones et al. approach, McGrath et al. argued that response bias indicators operate under the premise that they should either (a) moderate or (b) suppress the criterion-related validity of the substantive indicator (in this case, the self-report psychopathy measure) and that response distortion measures can detect response bias. According to McGrath et al., a stringent test of the response distortion hypothesis therefore requires evidence that the validity of self-report measures is compromised by response styles.

Baron and Kenny (1986) described moderation as reflecting the operation of a “variable that affects the direction and/or strength of the relation between an independent, or predictor, variable and a dependent, or criterion, variable” (p. 1174). Moderation can be represented as a statistical interaction between the predictor and a third variable that results in a change in the statistical prediction of the criterion. In contrast, suppression occurs when any third variable attenuates the relation between the predictor and the criterion. Whereas the third variable is highly associated with the predictor but negligibly associated with the criterion in classical suppression, the third variable is somewhat correlated with the criterion but more highly correlated with the predictor in other forms of suppression (i.e., net suppression). Thus, the statistical control of a suppressor variable results paradoxically in an increase in the correlation between the predictor and criterion (Cohen, Cohen, West, & Aiken, 2013).

To examine the hypothesis that response bias indicators statistically affect the criterion-related validity of a substantive indicator (heretofore referred to as the response bias hypothesis [RBH]), McGrath et al. (2010) identified 41 studies across various psychological domains (e.g., assessment of personality, emotional disorders, disability claims) that contained (a) self-report measures, (b) external criterion–criteria, and (c) response bias indicators, including measures of self-deception and impression management. They tested response bias indicators as both moderators and suppressors of the relations between self-report measures and external criteria and found weak evidence at best for both moderation and suppression. McGrath et al. reported some evidence of a biasing influence

from random or careless responding, but this conclusion was based on a small sample of studies and is not directly relevant to the hypotheses of this study. Although there are numerous ways to investigate the utility of response bias indicators (e.g., simulation studies involving instructional sets to “fake good” or “fake bad”), we concur with McGrath et al. that determining whether such indicators enhance the convergent validity of self-report psychopathy measures is the most rigorous approach to evaluating their validity. Specifically, the validity of the RBH requires consistent evidence that response distortion significantly moderates or suppresses the relation between self-reported psychopathy and external criteria.

## Psychopathy and Response Bias

Several prominent researchers in clinical psychology and allied fields have raised serious concerns regarding the use of self-report psychopathy measures, especially in forensic settings (e.g., Kelsey, Rogers, & Robinson, 2014). For example, Hart, Hare, and Forth (1996) asserted that “Behavioral checklists and self-report scales are poorly suited to assessing psychopathy because of their susceptibility to a variety of response biases” (p. 85). Similarly, Edens, Hart, Johnson, Johnson, and Olver (2000) contended that “Psychopathy is a difficult disorder to assess and diagnose, particularly through self-report questionnaires. [ . . . ] This is a major potential problem, because deceitfulness is construed as a core symptom of psychopathy” (p. 137). These quotations imply that psychopathic individuals are either unable or unwilling to present themselves accurately, ostensibly because they are prone to response distortions stemming from inflated self-concept, lack of insight, dishonesty, or all three. In addition to clinical speculation, research suggests that markedly psychopathic individuals are more prone to exaggerate or falsify their symptoms than are their less psychopathic counterparts (Gacono, Meloy, Sheppard, Speth, & Roske, 1995), even when there is no clear motivation to do so (Ekman, 1993).

Despite widespread skepticism of psychopathic individuals’ ability or willingness to report accurately on their traits, clinical suggestions to the contrary date back to writings at least as early as Cleckley (1941):

Although he will lie about any matter, under any circumstances, and often for no good reason, he may, on the contrary, sometimes own up to his errors (usually when detection is certain) and appear to be facing the consequences with singular honesty, fortitude, and manliness.” (pp. 341–342)

More recent research evidence demonstrates that psychopathic individuals tend to mangle only when there is clear incentive to do so (Rogers & Cruise, 2000). Furthermore, although psychopathy may be associated with malingering (cf. Niesten et al., 2015), evidence suggests that it is not associated with malingering success (Edens, Buffington, & Tomicic, 2000). Nevertheless, there is a clear gap between clinical lore and research evidence in the psychopathy literature regarding the RBH given that few studies have examined this hypothesis.

In an effort to address this gap, Ray et al. (2013) examined meta-analytically psychopathy’s relations with response bias indices, including indices of (a) underreporting, including both self-deception and impression management, and (b) overreporting or

“faking bad.” Contrary to most clinical lore, psychopathy scores were not positively associated with indices of response distortion in research studies. Psychopathy was either negatively or negligibly associated with underreporting, which was particularly true of F2 ( $r_w = -.16$ ). The findings for F1 depended on the measure: LSRP F1 was negatively associated with underreporting whereas PPI FD was essentially uncorrelated with underreporting ( $r_w = -.01$ ), perhaps owing to the largely maladaptive nature of LSRP F1 and the largely adaptive nature of PPI FD. Last, total psychopathy scores exhibited a small though significantly positive association with overreporting ( $r_w = -.11$ ). The Ray et al. findings suggest that self-report measures are not necessarily compromised by response bias, but they do not address the more direct question of whether such bias diminishes the validity of self-report psychopathy measures.

### Current Study

Although correlational analyses may yield some useful information regarding the potential effects of response bias on the validity of self-report psychopathy measures (Ray et al., 2013), the McGrath et al. (2010) methodology addresses this issue more directly. Nevertheless, only one published study has applied this methodology to psychopathy. Verschuere et al. (2014) applied the McGrath et al. methodology to psychopathy and examined impression management (i.e., underreporting) as a suppressor of the relations between self-reported psychopathy traits and self-reported externalizing behaviors (i.e., alcohol and drug abuse, indirect aggression, delinquency) among a sample of Belgian community members. Counter to the RBH, but broadly consistent with the findings of McCrae and Costa (1983), Verschuere et al. reported that controlling for underreporting weakened the association between psychopathy and external criteria. In other words, they found evidence for redundancy rather than suppression, suggesting that underreporting accounts for true variance in psychopathy. Nonetheless, Verschuere et al.’s study relied on self-reported external criteria and did not examine the potential effects of moderation.

In this study, we examined the statistical effects of response bias, particularly impression management, on the validity of self-report psychopathy measures in the statistical prediction of theoretically relevant external criteria in a large sample of ethnically diverse male and female offenders, including both incarcerated prisoners and nonincarcerated substance abusers referred for treatment. We elected to focus on the manifest psychopathy measures themselves, rather than on the latent variables reflecting their underlying constructs, given that we wished to ascertain whether the measures themselves, as routinely used in clinical practice and research, are adversely affected by response distortion. In addition, we examined not only total psychopathy scores but also scores on the major psychopathy subdimensions given the large body of evidence that these subdimensions often display pronounced differential correlates, including differential associations with normal-range personality traits and both internalizing and externalizing psychopathology (Derefinko & Lynam, 2006). In addition, we focused on external criteria that are nonquestionnaire based to minimize the potential role of method covariance.

Consistent with McGrath et al.’s (2010) suggestions and with findings in the broader personality literature, we predicted that

there would be little to no evidence for the RBH in our nonincarcerated sample. We in essence predicted the null hypothesis. That is, we did not expect to find consistent evidence of either moderation or suppression by response distortion measures, suggesting that response distortion does not diminish the validity of self-report psychopathy indices when there are no clear incentives for dissimulation or distortion. The positive response distortion measures used were designed largely to detect more overt overreporting (i.e., impression management). Although predicting the null hypothesis carries the risk of Type II errors, we used a large sample of individuals who were highly enriched with psychopathic traits, thereby ensuring that our design was adequately powered to detect even small statistical effects.

## Method

### Participants

Participants ( $N = 1,661$ ) were offenders who had been serving prison sentences (52%) or who had been court ordered to residential drug treatment programs (48%) at multiple sites in Florida (28%), Nevada (33%), Utah (19%), Oregon (13%), or Texas (7%; see Poythress et al., 2010, for a more detailed description of the sample and other measures administered). The sample largely comprised men (83%), and the mean age was 30.99 years ( $SD = 6.57$ ). In accord with the study design, which focused in part on race differences in psychopathy, only English-speaking African American (35%) and White (65%) individuals were enrolled; the sample comprised largely non-Hispanic individuals (87%). Self-reported educational attainment was as follows: high school not completed, 29%; general equivalency diploma, 27%; high school diploma, 21%; some college, 20%; college diploma, 2%; any graduate or professional school, .3%. Individuals currently taking psychotropic medications for active symptoms of psychosis (determined by self-report and chart review) were excluded, to minimize the potential influence of psychiatric symptoms on responding. In addition, offenders with a diagnosis of intellectual disability (formerly called mental retardation) or an IQ below 70 using a standardized instrument administered at the time of enrollment were excluded.

### Measures

**Psychopathic Personality Inventory (Lilienfeld & Andrews, 1996).** The PPI is a 187-item self-report inventory designed to assess the core psychopathic personality features, as opposed to overt antisocial or criminal behaviors. Participants answer each question using a 4-point Likert scale (1 = *false*, 2 = *mostly false*, 3 = *mostly true*, 4 = *true*). The PPI yields a total score; eight lower order subscale scores; scores on two higher order factors, PPI FD and PPI SCI (Benning et al., 2003); and three validity scales designed to detect biased or inconsistent responding: Variable Response Inconsistency (VRIN), Deviant Responding (DR), and Unlikely Virtues (UV). The VRIN scale, which consists of the sum of the absolute differences between 40 highly intercorrelated item pairs, aims to measure a respondent’s tendency to respond inconsistently to items of similar content. This study included DR and UV scales as response bias indicators (see the Response Bias Indices section). For this article, no participants were excluded on

the basis of deviant or inconsistent responding, as our purpose was to examine response biases and their effects on self-report psychopathy measures. Because of a small negative correlation between VRIN and UV scores,  $r = -.091$ ,  $p < .001$ , in this sample, excluding participants on the basis of high VRIN scores would slightly truncate the range of scores on UV, a key response bias indicator used in the analyses. Internal consistencies (alphas) ranged from .81 (PPI C) to .92 (PPI total).

**Levenson Self-Report Psychopathy Scale (Levenson et al., 1995).** The LSRP is a 26-item self-report measure, modeled largely after the items in the Psychopathy Checklist-Revised (Hare, 1991/2003), and designed to assess both the personality and behavioral features associated with psychopathy. Scores on two broad dimensions, reflecting primary ( $\alpha = .86$ ) and secondary psychopathy ( $\alpha = .73$ ), are provided.

**Personality Assessment Inventory–Antisocial Features Scale (PAI ANT; Morey, 1991).** The PAI is a multiscale self-report inventory assessing the features of clinical syndromes (11 scales), treatment-related characteristics (five scales), response bias (four scales), and interpersonal style (two scales). In this study, the ANT scale was used as a proxy of psychopathy given that the ANT scale was designed to assess both the personality and behavioral features relevant to antisocial behavior and psychopathy (Morey, 1991). For the purposes of this study, only total ANT scale scores were used (as opposed to subscale scores) to minimize the risk of Type I error ( $\alpha = .75$ ).

### Theoretically Relevant External Criteria

**Psychopathy Checklist–Revised (PCL-R; Hare, 1991/2003).** The PCL-R is the most widely used measure of psychopathy in forensic settings. It involves a clinical interview and review of institutional (file) records to assign ratings (0 = *not at all present* to 2 = *present*) on 20 features of psychopathy, yielding scores that can range from 0 to 40. In this study, global psychopathy scores (PCL-R total;  $\alpha = .82$ ), as well as Factor 1 (PCL-R F1;  $\alpha = .81$ ) and Factor 2 (PCL-R F2;  $\alpha = .68$ ) scores were obtained, per suggestions from the PCL-R manual and examined dimensionally. On the basis of 51 randomly selected cases, interrater reliability (ICC1) for the PCL-R total scores in this sample was .88 (Poythress et al., 2010). In addition, we extracted a one-item measure from the PCL-R of whether offenders had attempted suicide at any point in their lifetimes (see also Douglas et al., 2008) given F2's positive and F1's negative relations with suicide attempts (Verona, Patrick, & Joiner, 2001).

**Structured Clinical Interview for DSM-IV Axis II Personality Disorders–Antisocial Personality Disorder (SCID-II ASPD; First, Gibbon, Spitzer, Williams, & Benjamin, 1997).** The ASPD module of the SCID-II is a widely used semistructured psychiatric interview that assesses the DSM-IV and DSM-5 criteria for ASPD (American Psychiatric Association, 2013). In this study, a total symptom count variable was computed to assess ASPD dimensionally ( $\alpha = .83$ ). Previous studies using these data (Poythress et al., 2010) reported high interrater reliability for total symptom count (intraclass correlations [ICC<sub>1</sub>] = .86;  $n = 46$ ) and ASPD diagnoses ( $K = .74$ ;  $n = 50$ ).

**Go/no-go discrimination task (Newman & Kosson, 1986).** The go/no-go discrimination task is a widely used measure of passive-avoidance learning, when an individual learns to avoid

engaging in a behavior that results in punishment. Deficits in passive avoidance learning have long been associated with psychopathy (Lykken, 1957). During this task, participants must learn to press a button in response to the presentation of a particular two-digit stimulus (S+) to earn a monetary reward (e.g., 10 cents) and to suppress responding to other two-digit stimuli (S-). Following previous suggestions (Newman, Widom, & Nathan, 1985), the number of commission errors made during the second block of 40 trials (False Alarms Trial 2) was examined. Rather than competing for the actual cumulative 10-cent rewards, participants competed for a reward of \$25 to be given to the individual with the highest cumulative 10-cent reward total.

**Institutional infractions.** Research assistants obtained disciplinary reports for a 1-year period following recruitment into the study for offenders who were new admissions into their prison systems and coded these reports for infractions. Poythress et al. (2010) developed a coding scheme that consisted of nonaggressive infractions (e.g., possession of contraband), verbally aggressive infractions (e.g., spoken or written threats), or physically aggressive infractions (e.g., assault or battery with a deadly weapon), which were recoded to reflect a hierarchy of infraction types to reflect any (general) form of misconduct (including nonaggressive, verbally aggressive, or physically aggressive infractions), aggressive misconduct (verbal or physical aggression), and violent misconduct (physically aggressive infractions). Infractions were treated dimensionally as count variables. Descriptive statistics for these variables are included in Supplemental Table 1. Base rates for these three categories, reflecting when an infraction was committed at least once, were as follows: general infraction, 42.5%; verbally aggressive infraction, 27.4%; and physically aggressive or violent infraction, 6.2%.

### Response Bias Indices: Measures of Underreporting

**Unlikely Virtues scale (UV; Tellegen, 1982).** The UV scale consists of 14 items that describe implausibly virtuous behav-

Table 1  
Correlations of Self-Report Psychopathy Measures With Response Bias Indices

Measure	Underreporting			Overreporting		
	UV	PIM	DEF	DR	NIM	MAL
PPI T	<b>-.30</b>	<b>-.31</b>	<b>-.19</b>	<b>.24</b>	<b>.25</b>	<b>.15</b>
PPI FD	<b>.16</b>	<b>.22</b>	<b>.31</b>	<b>-.10</b>	<b>-.16</b>	<b>-.06</b>
PPI SCI	<b>-.47</b>	<b>-.56</b>	<b>-.46</b>	<b>.33</b>	<b>.46</b>	<b>.24</b>
PPI C	<b>-.01</b>	<b>.18</b>	<b>.10</b>	<b>.13</b>	<b>-.19</b>	<b>-.09</b>
LSRP F1	<b>-.22</b>	<b>-.29</b>	<b>-.18</b>	<b>.27</b>	<b>.29</b>	<b>.21</b>
LSRP F2	<b>-.38</b>	<b>-.58</b>	<b>-.53</b>	<b>.28</b>	<b>.47</b>	<b>.22</b>
PAI Ant	<b>-.38</b>	<b>-.51</b>	<b>-.36</b>	<b>.17</b>	<b>.39</b>	<b>.21</b>

*Note.* Significant correlations are italicized ( $p < .05$ ) or in boldface type ( $p < .001$ ) for emphasis. PPI = Psychopathic Personality Inventory; T = total; FD = Fearless Dominance; SCI = Self-Centered Impulsivity; C = Coldheartedness; LSRP = Levenson Self-Report Psychopathy Scale; F1 = Factor 1; F2 = Factor 2; PAI = Personality Assessment Inventory; Ant = Antisocial Features Scale; PCL-R = Psychopathy Checklist-Revised; UV = unlikely virtues; PIM = positive impression management; DEF = defensiveness; DR = deviant responding; NIM = negative impression management; MAL = malingering.

iors based on the seven deadly sins (e.g., gluttony, wrath) and appears to primarily index impression management. These items were embedded within the PPI with permission from the MPQ's author.

**PAI Positive Impression Management (PIM; Morey, 1991).** PIM, one of four primary validity scales embedded within the PAI, is a nine-item measure of impression management that assesses the endorsement of favorable characteristics and the denial of minor shortcomings.

**PAI Defensiveness (PAI DEF; Morey, 1991).** DEF comprises eight configural features that distinguish the PAI profiles of individuals instructed to present themselves positively from the profiles of community or clinical participants. High scores on DEF are intended to reflect overt defensiveness, although some research has called its utility into question.

### Response Bias Indices: Measures of Overreporting

**PPI Deviant Responding (PPI DR; Lilienfeld & Andrews, 1996).** DR is a 10-item scale embedded within the PPI that comprises items measuring superficially plausible, yet nonexistent, psychological symptoms. The DR aims to detect malingering or other "faking bad" response sets or styles, as well as otherwise bizarre responses.

**PAI Negative Impression Management (PAI NIM; Morey, 1991).** NIM is a nine-item scale comprising seemingly pathological items with relatively low endorsement rates in both nonclinical and clinical samples as well as a negativistic response style that is potentially different from malingering. These items were constructed to depict valid psychopathological symptoms that are rare or nonexistent, even among clinical populations.

**PAI Malingering (PAI MAL; Morey, 1991).** The malingering scale comprises eight configural features of the PAI profile that appear more frequently in the profiles of individuals instructed to feign mental disorder, particularly compared with the profiles of respondents with severe mental illness. The number of positive features is summed, creating a score from 0 to 8. Each feature of the MAL scale is generally observed with far greater frequency in feigning samples, as opposed to clinical or community samples.

### Procedure

Participants completed a battery of measures relevant to general personality, psychopathy and antisocial behavior, including interview and self-report measures of psychopathy and antisocial personality disorder (ASPD), and laboratory tasks associated with psychopathy. Research assistants also obtained disciplinary reports for a 1-year period following recruitment into the study for those newly admitted to prison. The clinical interviewers, who were doctoral students in clinical psychology, underwent a formal 2-day training in the PCL-R by an expert in the measure as well as a formal training in the SCID ASPD module by the second author.

### Statistical Analysis

We used regression models to address the potential statistical effects of response bias indicators on the relations between self-

report psychopathy measures and theoretically relevant external criteria by means of statistical moderation and suppression. We used statistical procedures outlined by Baron and Kenny (1986) and adapted by Preacher and Hayes (2004). In the case of certain external criteria (i.e., infraction data) that were severely positively skewed, with a preponderance of zeros, we used zero-inflated Poisson (ZIP) regression models.

**Statistical moderation.** We used the Preacher and Hayes' (2004) PROCESS macro for SPSS to estimate simple moderation models. PROCESS calculates the product of X and M (the moderator), mean centers all predictor variables, and calculates the proportion of variance in Y uniquely attributable to the moderation of X's effect by M. Bootstrapped confidence intervals (derived from 1000 bootstrapped samples) were interpreted to determine whether a significant statistical interaction was present in each moderation model.  $R^2$  change due to the interaction is presented for each moderation model.

Because we were in essence testing the null hypothesis, we did not correct for family-wise error, as we intended to provide as adequately powered a set of statistical tests as possible. PROCESS computes the Johnson-Neyman values for all continuous variables. It derives the values of the moderator such that the ratio of the conditional effect to the standard error is equal to the critical  $t$  value associated with  $p = \alpha$ , where  $\alpha$  is the chosen significance level (.05, in this case). We interpreted Johnson-Neyman values to determine whether the interaction effect was significant either above or below the mean level of the moderator, similar to a simple slopes approach.

**Statistical suppression.** Suppression was examined by means of mediation modules to estimate the total, direct, and indirect effects for each model. Instances of suppression are evident in mediation models when the direct effect (X on Y) is greater than the indirect effect (X \* M on Y). In this case the addition of the suppressor variable, M (here, the response bias indicator), should attenuate the prediction of Y (the external criterion). Within the context of the RBH, instances of suppression would suggest that controlling for the response bias indicator in the prediction of the external criterion should increase the relation between the self-report psychopathy indicator and the external criterion. Thus, response bias would attenuate the relation between the self-report criterion and the more "objective" criterion.

As recommended by experts on mediation (Preacher & Hayes, 2004, 2008), we estimated the total, direct, and indirect effects of response bias indices on the relations between self-report psychopathy measures and theoretically relevant external criteria using 1000 bias-corrected bootstrap samples. The bias-corrected bootstrap does not assume that the indirect effect is normally distributed and, thus, is one of the most statistically powerful mediation tests available (Fritz & MacKinnon, 2008). Along with bootstrapped confidence intervals, two measures of effect size were calculated for each analysis containing a continuous outcome variable (Y; in this case, the theoretically relevant external criterion); effect size estimates were not available for dichotomous indicators (i.e., institutional infractions, PCL-R suicide attempts).

To address the need for a meaningful effect size estimate for indirect models, Preacher and Kelley (2011) developed several effect size estimates that better apply to the indirect effect, ostensibly the most important information gleaned from a mediation analysis. For the purposes of this manuscript, we focus

on two of available estimates,  $R^2$  and Kappa-squared ( $K^2$ ), the latter of which Preacher and Kelley prefer. Nevertheless, given the widespread use of  $R^2$  estimates in moderation and mediation analyses, we provided them in this article. The use of  $R^2$  estimates of indirect effects comes with several limitations, namely that  $R^2$  is not a true expression of the proportion of variance explained (i.e.,  $R^2$  is negative in the case of suppression) and that it is not a proportion of variance explained by the actual maximum amount of variance possible. Given the distributional properties and unreliability of most measures, it is unlikely that the maximum amount of variance explained by a given model is actually 100%, although this is often assumed erroneously. Thus, where  $R^2$  values may appear small, they may actually be fairly large when considering the maximum possible effect (when taking measurement error and skewness into account). To address the issues inherent to  $R^2$ , Preacher and Kelley developed an effect size estimate that reflects the magnitude of the indirect effect relative to the maximum possible indirect effect.  $K^2 = 0$  reflects no linear indirect effect, and  $K^2 = 1$  reflects that the indirect effect is the maximum possible value.

### Simulation Analyses

To account for the problems associated with multiple testing and ascertain further whether the significant moderation and suppression findings were potentially attributable to Type I error, we used the programming language R to conduct a simulation analyses. We used the simulation to determine the number of significant moderation and suppression effects that would occur by chance when the true relation between the moderator-suppressor and the dependent variable was zero. First, we randomly simulated each hypothetical self-report measure score from a standard normal probability distribution. Then, we used the estimated sample correlations as the parameter values defining the correlations among our simulated psychopathy and response bias variables to ensure the relationships between our simulated psychopathy and response bias variables resembled their observed relationships. For example, the correlation between the PPI total self-report measure scores and the PCL-R total criterion scores was .42 (see Table 1), so the simulation value defining the effect was set to .42, and the effect of every response bias index on the PCL-R total scores was set to zero. Last, the residual terms used for both the response bias indices and the criterion variables were simulated from a random normal distribution with a mean of zero and a variance equal to one minus the squared effect of the self-report measure to ensure that the total variance for each of the response bias indices and criterion variables would be equal to one. We simulated each regression model (54 models for a single independent variable) 1,000 times to ensure the stability of our results.

## Results

### Zero-Order Correlations Between Self-Report Psychopathy Measures and Response Bias Indices

Table 1 presents the correlations between self-report psychopathy measures and response bias indices. Overall, correlations

were consistent with previous literature (Ray et al., 2013). With the exception of PPI FD and PPI C, psychopathy measures were significantly negatively associated with indices of underreporting. The correlations for PPI SCI and LSRP F2 were larger in magnitude and were medium to large in magnitude, consistent with the notion that they reflect more maladaptive aspects of psychopathy than do F1 scores.

PPI FD was positively associated with all indicators of underreporting, although the effects were generally small to medium in magnitude. PPI C was slightly positively associated with two indicators of underreporting (i.e., PIM, DEF) and was unassociated with UV, although the latter two correlations were small in magnitude. Also consistent with the Ray et al. (2013) meta-analysis, the psychopathy measures with the exception of PPI FD and PPI C were positively associated with indices of overreporting. PPI FD was slightly significantly negatively associated with overreporting measures. PPI C was slightly positively associated with DR, but was slightly negatively associated with NIM and MAL. (Descriptive statistics and intercorrelations among of psychopathy and ASPD measures and response bias indicators are presented in the online supplementary Tables 1 and 2, respectively. Relations between these variables were as expected and consistent with previous literature.)

### Moderation Analyses

Because of the large number of analyses ( $N = 378$ ), each regression is not presented here (see Table 2 for a summary of the results from the moderation analyses). Rather, overall findings subdivided by self-report psychopathy measure are presented. In sum, 26 of the 378 (7%) moderation analyses conducted were statistically significant at an alpha level of .05. For analyses in support of the RBH (i.e., significant effects at high levels of the response bias indicator), 12 of the 378 (3%) of the analyses conducted were significant. Of the 12 analyses that were significant in support of the response bias hypothesis, 0 analyses for the following psychopathy indicators were significant in the direction of the RBH: PPI FD, PPI SCI, and LSRP F2. Three of the 54 (6%) PPI total analyses ( $\Delta R^2$  ranged from .00 to .00) and LSRP F1 ( $\Delta R^2$ : .00 to .01), 2 of 54 analyses (4%) PPI C analyses ( $\Delta R^2$ : .00 to .01), and 4 (7%) of the PAI ANT analyses ( $\Delta R^2$ : .00 to .06) were significant in the direction of the response bias hypothesis. Self-report psychopathy measures with a potentially above-chance level (>5% of analyses, not accounting for the correlations among the dependent measures) number of significant moderation analyses in the direction of the RBH were explored further for patterns of theoretical interest. There were no discernable patterns in terms of the statistically significant analyses in support of the RBH.

**Moderation simulations.** The results of this analysis yielded significant findings about 5% of the time, across all seven independent variables. Given that only 3% of our analyses supported the RBH, these simulation analyses suggest that our positive findings are plausibly attributable to chance.

### Suppression Analyses

Table 3 presents a summary of the suppression analyses. Again, because of the large number of analyses ( $N = 378$ ), each regres-



Table 2  
Summary of Moderation Analyses

Self-report measure	No. of analyses	Moderation		In support of RBH		$R^2$ increase range for results in direction of RBH	
		<i>N</i> sig	% sig	<i>N</i> sig	% sig		
PPI T	54	5	9.26	3	5.56	.00	.00
PPI FD	54	0	0.00	0	0.00	—	—
PPI SCI	54	2	3.70	0	0.00	—	—
PPI C	54	6	11.11	2	3.70	.00	.01
LSRP F1	54	5	9.26	3	5.56	.00	.01
LSRP F2	54	1	1.85	0	0.00	—	—
PAI Ant	54	7	12.96	4	7.41	.00	.06
Total	378	26	6.88	12	3.17	.00	.06

Note. RBH = response bias hypothesis; sig = significant; PPI = Psychopathic Personality Inventory; T = total; FD = fearless dominance; SCI = self-centered impulsivity; C = Coldheartedness; LSRP = Levenson Self-Report Psychopathy Scale; F1 = Factor 1; F2 = Factor 2; PAI = Personality Assessment Inventory; Ant = Antisocial Features Scale.

sion is not presented here. Instances of theoretically relevant findings in support of the response bias hypothesis are addressed in greater detail.  $R^2$  and  $K^2$  values were reported for all significant suppression analyses containing continuous (but not dichotomous) outcome variables. Overall, there was somewhat more support for suppression than for moderation, although we also found consistent evidence for redundancy in many cases (see also Verschuere et al., 2014). One hundred ninety-three of the 378 (51%) analyses derived from the mediation modules analyzed were statistically significant. Of these 193 significant analyses, precisely one third (17.2% of the total number of analyses) were significant in the direction of suppression (i.e., in the direction of the RBH). Of the 54 analyses from the mediation modules analyzed for both LSRP F1 and LSRP F2, fewer than 5% of the analyses were significant for suppression (4% significant for each indicator). Thus, patterns of significance will not be discussed here, as these significant results are plausibly attributable to Type I error.

For PPI Tital scores, 5 out of 54 analyses were significant for suppression ( $\Delta R^2$  values ranged from  $-.01$  to  $.01$ ; average  $K^2$  value =  $.02$ ) but were associated with very small effect sizes. Four

of the five significant analyses were for PCL-R F1 scores, particularly when an overreporting index (i.e., UV, PIM, DEF) indicator was used. For PPI SCI analyses, eight of the 54 analyses were statistically significant ( $\Delta R^2$ :  $-.03$  to  $.03$ ; average  $K^2$  =  $.05$ ). Of the eight significant analyses, four were for PCL-R F1 scores and three were for PCL-R total scores, particularly in the case of analyses using underreporting mediators (all but one of the eight significant analyses were with an underreporting mediator; see Table 4 for a summary of PPI SCI analyses predicting PCL-R F1 and total scores). Aside from these two patterns, the remaining significant analyses appear to be largely random.

For PPI FD scores, 25 of the 54 analyses were significant for suppression ( $\Delta R^2$ :  $-.03$  to  $.00$ ; average  $K^2$  =  $.03$ ), indicating that the average indirect effect explains approximately 3% of the maximum possible indirect effect. In analyses in which FD was associated with more maladaptive features of psychopathy (i.e., PCL-R F2, PCL-R total, SCID ASPD), each of the six mediation analyses for each criterion variable was significant (see Table 4 for a summary of PPI FD analyses predicting PCL-R F2, PCL-R total,

Table 3  
Summary of Suppression Analyses

Self-report measure	Regression analyses							Simulation analyses				
	<i>N</i> analyses	Mediation		Suppression		Effect sizes for results in direction of RBH		<i>N</i> analyses	Mediation		Suppression	
		<i>N</i> sig	% sig	<i>N</i> sig	% sig	Average $K^2$	$R^2$ range		<i>N</i> sig	% sig	<i>N</i> sig	% sig
PPI T	54	16	29.63	5	10.42	.026	$-.01$ to $.01$	54,000	1,313	2.43	1,347	2.50
PPI FD	54	24	44.44	25	46.30	.043	$-.03$ to $.00$	54,000	1,378	2.55	1,336	2.47
PPI SCI	54	13	24.07	8	14.81	.055	$-.03$ to $.03$	54,000	1,395	2.58	1,399	2.59
PPI C	54	23	42.59	17	31.48	.021	$-.01$ to $.00$	54,000	1,356	2.51	1,378	2.55
LSRP F1	54	16	29.63	2	3.70	.024	$-.01$ to $.01$	54,000	1,298	2.40	1,301	2.41
LSRP F2	54	19	35.19	2	3.70	.050	$-.01$ to $.00$	54,000	1,343	2.49	1,306	2.42
PAI Ant	54	17	31.48	6	11.11	.036	$-.02$ to $.00$	54,000	1,312	2.43	1,330	2.46
Total	378	128	33.86	65	17.20	.032	$-.03$ to $.03$	54,000	9,395	2.49	9,397	2.49

Note. RBH = response bias hypothesis; sig = significant; PPI = Psychopathic Personality Inventory; T = total; FD = fearless dominance; SCI = self-centered impulsivity; C = coldheartedness; LSRP = Levenson Self-Report Psychopathy Scale; F1 = Factor 1; F2 = Factor 2 Psychopathy; PAI = Personality Assessment Inventory; Ant = Antisocial Features Scale.

Table 4  
Notable Suppression Findings

Predictor	Mediator	Criterion	Significant indirect effect?	K <sup>2</sup>	R <sup>2</sup>
PPI FD	UV	SCID ASPD	Yes	.04	-.01
	PIM		Yes	.07	-.02
	DEF		Yes	.09	-.02
	DR		Yes	.01	.00
	NIM		Yes	.04	-.01
	MAL		Yes	.01	.00
PPI FD	UV	PCL-R F2	Yes	.03	-.01
	PIM		Yes	.05	-.02
	DEF		Yes	.08	-.03
	DR		Yes	.02	-.01
	NIM		Yes	.04	-.01
	MAL		Yes	.00	-.00
PPI FD	UV	PCL-R Total	Yes	.02	-.01
	PIM		Yes	.03	-.01
	DEF		Yes	.04	-.02
	DR		Yes	.01	-.01
	NIM		Yes	.03	-.01
	MAL		Yes	.00	.00
PPI SCI	UV	PCL-R F1	Yes	.04	-.01
	PIM		Yes	.08	-.02
	DR		No		
	NIM		Yes	.03	.00
	MAL		No		
PPI SCI	UV	PCL-R Total	Yes	.03	.01
	PIM		Yes	.07	-.01
	DEF		Yes	.06	-.01
	DR		No		
	NIM		No		
	MAL		No		
PPI C	UV	PCL-R F2	No		
	PIM		Yes	.039	-.01
	DEF		Yes	.02	-.01
	DR		No		
	NIM		Yes	.04	-.01
	MAL		Yes	.01	.00
PPI C	UV	PCL-R Total	No		
	PIM		Yes	.02	-.01
	DEF		Yes	.01	.00
	DR		No		
	NIM		Yes	.03	-.01
	MAL		Yes	.01	.00

Note. PPI = Psychopathic Personality Inventory; FD = fearless dominance; F1 = Factor 1 Psychopathy; F2 = Factor 2 Psychopathy; PCL-R = Psychopathy Checklist-Revised; SCID ASPD = Structured Clinical Interview for the *Diagnostic and Statistical Manual of Mental Disorders*, Antisocial Personality Disorder; UV = unlikely virtues; PIM = positive impression management; DEF = defensiveness; DR = deviant responding; NIM = negative impression management; MAL = malingering.

and SCID ASPD scores). The remaining significant analyses do not follow any evident pattern.

For PPI C, 17 of the 54 analyses were statistically significant for suppression ( $\Delta R^2$  values:  $-.01$  to  $.00$ ; average  $K^2$  value =  $.02$ ). There were several patterns detected, particularly when PPI C was statistically associated with more maladaptive aspects of psychopathy (i.e., PCL-R F2 and total scores). Of the 17 significant analyses, four were for PCL-R F2 scores and five were for PCL-R total scores (see Table 4 for a summary of the PPI C analyses predicting PCL-R F2 and total scores). A mix of both underreporting and overreporting response bias indicators was significant in

these analyses. In the case of PAI ANT scores, six of the 54 analyses conducted were significant for suppression ( $\Delta R^2$  values:  $-.02$  to  $.00$ ; average  $K^2$  value =  $.02$ ). Three of the six analyses were significant when PAI ANT predicted PCL-R F1 scores with an underreporting mediator. No other clear patterns in significance arose.<sup>1</sup>

**Suppression simulations.** The results of this simulation can be found in Table 3. Of the 378,000 models that were simulated, 4.98% of them estimated a significant effect of the response bias indicator on the criterion variable despite the true effect of zero. Moreover, approximately 50% of the significant effects were in support of the RBH and 50% were against the RBH. Specifically, the total number of significant mediation analyses was 9,395 (2.49% of every estimated effect), and the total number of significant suppression analyses was 9,397 (2.49% of every estimated effect), which suggests virtually equal evidence for and against the RBH. Thus, 5% of the estimated effects were significant by chance and of those significant effects, half would support the RBH and half would not. Given that 17.2% of our analyses supported the RBH, we can conclude that the number of suppression findings exceeded chance levels.

## Discussion

We sought to examine the statistical influence of response distortion on the validity of self-report psychopathy measures in the statistical prediction of theoretically relevant external criteria. In light of the McGrath et al. (2010) findings (see also McCrae & Costa, 1983) of negligible evidence that response biases moderate or suppress the relations between self-report measures and theoretically relevant external indicators outside of the self-report domain, we predicted that there would be little to no evidence for the RBH in a large dataset of *nonincentivized* offenders. To do so, we conducted over 750 moderation and suppression analyses in a large, ethnically diverse sample of male and female offenders, including both prisoners and substance abusers, to examine this theoretically and practically significant hypothesis.

## Summary of Results

Overall, with two potential exceptions of unclear practical significance (see next section), our findings offer little consistent support for the RBH, which suggests that psychopathic individuals are unwilling or unable to provide valid responses on self-report indicators. We found that a mere 3% of the 504 moderation analyses and conducted were statistically significant and in support of the RBH. On examination of the significant moderation analyses, no clear pattern emerged. The results of the moderation simulation analysis yielded significant findings about 5% of the time across all seven dependent variables, which outnumbered the 3% of the significant moderation analyses. Thus, our significant findings are plausibly due to Type I error.

<sup>1</sup> Supplementary moderation and suppression analyses removing all female participants revealed results extremely similar to those presented here. Full analyses can be obtained from the first author on request.

## Notable Exceptions

We found that response bias, particularly underreporting, often suppressed the relations between largely adaptive aspects of psychopathy and the largely maladaptive features or correlates of psychopathy (i.e., PCL-R F2, PCL-R total, and SCID ASPD). Suppression was most evident for PPI FD and C, suggesting that the residual variance of PPI FD and C is more closely aligned with maladaptive aspects of psychopathy after removing the adaptive variance present in impression management indices. Specifically, 17.2% of suppression analyses were significant in the direction of the RBH. Simulation analyses yielded significant findings in support of the RBH only 2.5% of the time, suggesting that the significant suppression findings are unlikely to be attributable to Type I error. Nevertheless, even these effects tended to be small in magnitude, accounting for anywhere from 1% to 4% of the variance. Thus, although potentially theoretically interesting and meaningful, these findings may be of little practical significance in (nonincentivized) assessment contexts. Moreover, these suppression findings were limited to interview-based measures (PCL-R, ASPD module), suggesting that further examination of the generalizability of our suppression findings is necessary.

These significant suppression findings were specific to PPI FD and C, the former of which has sometimes been criticized for being a primarily adaptive feature of psychopathy (Miller & Lynam, 2012; but see Lilienfeld et al., 2012, for a response). In terms of their relations with general personality variables, FD and C are largely associated with low neuroticism and high levels of emotional resilience, which are presumably adaptive personality traits (Lilienfeld et al., 2012). Similarly, underreporting, or defensiveness, is largely associated with low neuroticism and high levels of resilience and self-esteem (Ones et al., 1996). It is therefore plausible that when controlling statistically for underreporting (a somewhat adaptive component of response bias; Block & Thomas, 1955), largely but not entirely adaptive indicators of psychopathy (e.g., PPI FD) become more closely aligned with maladaptive indicators (e.g., PCL-R F2).

Nevertheless, even our statistically significant suppression analyses do not necessarily provide evidence for the RBH. Response distortion indices, including measures of positive impression management, are imbued with genuine personality variance. High scores on these indices are associated with low levels of neuroticism or emotional stability and to a lesser degree, high levels of extraversion, agreeableness, and openness from a five-factor model perspective (McCrae & Costa, 1983), as well as high levels of honesty–humility from the perspective of the HEXACO model of personality (de Vries, Zettler, & Hilbig, 2014). As a consequence, removing adaptive variance from more adaptive aspects of psychopathy (e.g., PPI FD, PPI C) may yield a residual variable that is tied predominantly to pathological indices (e.g., Factor 2 psychopathy, ASPD). A more rigorous test of this possibility would necessitate analyses in which measures of neuroticism, agreeableness, and honesty–humility are statistically controlled in the relation between largely adaptive measures of psychopathy and maladaptive indices.

In exploratory analyses not presented here, we conducted suppression analyses in which self-reported psychopathy scores statistically predicted predominantly maladaptive indices (e.g.,

PCL-R F2, SCID ASPD) while covarying for a measure of Negative Emotionality, which is highly related to neuroticism. Specifically, we factor analyzed the PAI scales to develop a composite measure of Negative Emotionality (PAI NEM).<sup>2</sup> The results were largely nonsignificant, suggesting that underreporting does not predict above and beyond negative emotionality. Similarly, partial correlations between self-report psychopathy measures and external criteria became largely significant and substantially larger in magnitude after covarying for PAI NEM. We also conducted exploratory analyses to ascertain whether similar suppression results emerged with normal-range personality variables, in this case PAI Dominance (PAI DOM) and Warmth (PAI WRM), the latter of which appears to be a good proxy for Agreeableness (Hopwood et al., 2011). Partial correlations between self-report psychopathy measures and external criteria were largely significant and higher in magnitude after covarying for PAI DOM and WRM. Taken together, these significant suppressor effects buttress the contention that our positive suppression findings are because impression management indices are imbued with genuine personality variance, rather than to the problematic validity of self-report psychopathy measures.

## Limitations and Future Directions

Our findings come with several limitations and leave other questions unanswered. First, one potential criticism of this study is the potential for Type I error arising from the large number of analyses conducted. Although we considered a post hoc alpha correction, this procedure would have biased our analyses in favor of the null hypothesis and thereby provided a less stringent test of the RBH. Had we used a post hoc adjusted alpha for the 756 analyses we conducted, the alpha would have decreased to .000066 (.05/756), rendering the vast majority of analyses nonsignificant. Instead, a more conventional alpha level revealed several theoretically meaningful patterns of significance, particularly with predominantly adaptive indicators derived from the PPI. If anything, our lenient alpha level gave the RBH more than an ample opportunity for corroboration. Nevertheless, with the potential exceptions of questionable practical significance already observed, the RBH was not supported by the evidence presented here.

Second, our results may be limited to individuals who are largely “unsuccessful” given that we examined participants in prison or court-mandated substance treatment settings. Our findings may not extend to more “successful” psychopathic individuals, including (a) criminal individuals who have escaped detection by the legal system and (b) individuals who have channeled their

<sup>2</sup> A negative emotionality (NEM) composite was estimated using exploratory factor analysis (EFA) of all Personality Assessment Inventory (PAI) scales excluding the validity scales and the Antisocial (ANT) and Aggression scales (which clearly overlapped conceptually with the external criteria). The scree plot pointed to a clear one factor solution. After forcing the EFA to extract one factor, which we interpreted as NEM, the PAI scales with the highest loadings (>.70) on this factor were Anxiety, Anxiety Related Disorders, Depression, Schizophrenia, and Borderline. PAI Somatic Complaints, Paranoia, Suicide, Stress, Mania, Alcohol Problems, and Drug Problems exhibited loadings ranging from .3 to .7. We derived the NEM composite by summing participants' scores across the scales on this factor, weighted by variables' loadings. Factor loadings are available from the first author on request.

psychopathic tendencies into constructive outlets, such as business, law enforcement, and military combat (see Smith, Watts, & Lilienfeld, 2014; and Widom, 1977, for discussions).

Third, our participants were nonincentivized. Because the data were confidential and used exclusively for research purposes, the results of data collection had no effect on individuals' standing for parole, standing within their alcohol–drug treatment program, treatment prospects, and so on. Hence, our findings may not generalize to an incentivized sample, such as (a) individuals in litigation; (b) offenders being evaluated for parole, competency, or sanity; or (c) parents undergoing custody evaluations (see Edens & Ruiz, 2006). It is unclear whether psychopathic individuals are more successful malingerers than are nonpsychopathic individuals, although the admittedly limited evidence suggests that they are not (Poythress et al., 2010); however, the former are certainly more willing than the latter to malingering. Thus, in incentivized settings, psychopathic individuals may be especially prone to response bias (Kelsey, Rogers, & Robinson, 2014). Nevertheless, at least within our nonincentivized sample, the validity of self-report psychopathy measures was not especially adversely affected by response bias. Therefore, the McGrath et al. (2010) challenge may be difficult to meet within modal research settings, in which participants have little to no obvious reason to engage in aberrant responding. Hence, we encourage researchers to examine the extent to which our findings extend to real-world settings.

Fourth, Morey (2012) raised several useful challenges to the McGrath et al. (2010) meta-analysis, one of which was insufficient power to detect statistical moderation and suppression in cases in which the predictor and criterion variables are weakly or negligibly correlated. In this study, several criterion variables exhibited low to medium correlations with the predictors. For example, the self-report psychopathy measures correlated negligibly with passive avoidance errors on the go/no-go task. In addition, the maximum correlation between self-report psychopathy measures and ASPD symptom counts was only  $r = .15$  (with PPI SCI). The low correlations between the criterion and predictor variables may have impeded our ability to detect certain instances of moderation and suppression.

Fifth, we examined only the potential effects of response bias on manifest psychopathy measures themselves rather than on latent variables ostensibly reflecting the constructs assessed by these measures. We did so to address the pragmatically important question of the extent to which the validity of these measures as used in clinical practice and research is adversely affected by response distortion. In future research, it may be of interest to examine to place these measures within a latent variable model, which would allow investigators our central questions free of measurement error. Finally, as noted by other authors (e.g., de Vries et al., 2014), most positive response distortion measures are a complex mix of self-deception and impression management. Although the scales we used were designed largely to detect impression management, the extent to which our results extend to subtler forms of positive response distortion (i.e., self-deception) requires further investigation.

Our findings leave several important questions unresolved. First, our findings do not address the question of whether the statistical effects of moderation are most pronounced at high, as opposed to moderate, levels of response bias. Subsidiary analyses, again not presented here, in which the squared term for each response bias indicator was entered following its main effects, yielded negligible support for this possibility. Nevertheless, these

analyses may have been underpowered given that the power to detect significant curvilinear effects of moderation is small. Hence, additional research examining the possibility of curvilinear effects in even larger samples is necessary.

Second, our results leave resolved the question of whether and when elevated scores on validity scales, such as measures of impression management, should lead clinicians to call into question the self-reported psychopathy scores of individual clients in clinical practice. This question harkens back to the nomothetic-idiographic distinction in personality, and the accompanying clinical versus actuarial debate (Grove & Meehl, 1996). As Meehl (1954) observed in his classic book, most of the questions addressed in clinical psychology, including that examined in our article, are nomothetic (“What do scale elevations tell us about individuals in general?”), whereas the question confronting the everyday clinician is idiographic (“What does this scale elevation tell me about my client?”). Meehl (1957) further noted that unless there is a compelling reason not to do so, we should generally “go with the formula” rather than “use our heads.” Nevertheless, as he acknowledged, the question of whether and when actuarial formulas should be overruled by rare exceptions (“broken leg cases”) is not entirely resolved. In our view, the most scientifically defensible interpretation of our findings is that, pending the potential discovery of moderators that qualify our findings for identifiable subsets of individuals, the default position should be that a given client's score on a self-report psychopathy scale should not be automatically dismissed on the grounds of high scores on an impression management scale. Nevertheless, this conclusion holds only for clients for whom therapists have no clear-cut reason to anticipate incentives for response distortion.

Future studies would benefit from the inclusion of other theoretically relevant indicators not available in the current study, such as other interview measures of psychopathy (e.g., the Comprehensive Assessment of Psychopathic Personality; Cooke, Hart, & Logan, 2005), additional laboratory tasks relevant to psychopathy, and informant reports of psychopathy. Further examination of external criteria within psychopathy's nomological network will be necessary to conclude that self-report measures are not adversely affected by response bias. In particular, informant reports provide a rich source of personality and personality disorder data (Vazire, 2006). In addition, they allow for aggregation across multiple observers (Block, 1961), thereby potentially reducing measurement error. Informant reports tend to be potent predictors of behavior, although they correlate at least moderately with self-reports (Miller, Jones, & Lynam, 2011). Indeed, they often provide incremental validity above and beyond self-reports for relevant criteria (Connelly & Ones, 2010), including externalizing behaviors (Jones & Miller, 2012) and personality traits (e.g., extraversion; John & Robins, 1993).

## Conclusions

With the potential exception of the largely adaptive features of psychopathy, which are controversial insofar as their relevance to the condition (e.g., Miller & Lynam, 2012), the current results offer little consistent evidence for the RBH in a nonincentivized context with respect to the assessment of psychopathy. In general, response distortion, including impression

management, does not usefully moderate or suppress of self-reported psychopathy in statistically predicting outcomes (see also Verschuere et al., 2014). Hence, our findings may allay researchers' concerns that a tendency to make oneself look unrealistically healthy or unhealthy necessarily compromises the validity of self-report psychopathy measures in research settings. Finally, our findings suggest that impression management scales in isolation should not be used in research to discard or adjust self-report psychopathy scores.

## References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental Disorders* (5th ed.). Washington, DC: Author.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182. <http://dx.doi.org/10.1037/0022-3514.51.6.1173>
- Benning, S. D., Patrick, C. J., Hicks, B. M., Blonigen, D. M., & Krueger, R. F. (2003). Factor structure of the psychopathic personality inventory: Validity and implications for clinical assessment. *Psychological Assessment*, *15*, 340–350. <http://dx.doi.org/10.1037/1040-3590.15.3.340>
- Ben-Porath, Y. S., & Waller, N. G. (1992). "Normal" personality inventories in clinical assessment: General requirements and the potential for using the NEO Personality Inventory. *Psychological Assessment*, *4*, 14–19. <http://dx.doi.org/10.1037/1040-3590.4.1.14>
- Block, J. (1961). *The Q-sort method in personality assessment and psychiatric research* (No. 457). Springfield, IL: Thomas. <http://dx.doi.org/10.1037/13141-000>
- Block, J., & Thomas, H. (1955). Is satisfaction with self a measure of adjustment? *Journal of Abnormal and Social Psychology*, *51*, 254–259. <http://dx.doi.org/10.1037/h0048246>
- Cleckley, H. (1941). *The mask of sanity: An attempt to reinterpret the so-called psychopathic personality*. St. Louis, MO: Mosby.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis*. Mahwah, NJ: Erlbaum.
- Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, *136*, 1092–1122. <http://dx.doi.org/10.1037/a0021212>
- Cooke, D., Hart, S., & Logan, C. (2005). Comprehensive Assessment of Psychopathic Personality Disorder: Institutional Rating Scale, Version 1.1. [Unpublished manual].
- Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment*, *13*, 171–188. <http://dx.doi.org/10.1037/1040-3590.13.2.171>
- Davis, C. G., Thake, J., & Weekes, J. R. (2012). Impression managers: Nice guys or serious criminals? *Journal of Research in Personality*, *46*, 26–31. <http://dx.doi.org/10.1016/j.jrp.2011.11.001>
- Derefinko, K. J., & Lynam, D. R. (2006). Convergence and divergence among self-report psychopathy measures: A personality-based approach. *Journal of Personality Disorders*, *20*, 261–280. <http://dx.doi.org/10.1521/pedi.2006.20.3.261>
- de Vries, R. E., Zettler, I., & Hilbig, B. E. (2014). Rethinking trait conceptions of Social Desirability Scales: Impression management as an expression of honesty-humility. *Assessment*, *21*, 286–299.
- Douglas, K. S., Lilienfeld, S. O., Skeem, J. L., Poythress, N. G., Edens, J. F., & Patrick, C. J. (2008). Relation of antisocial and psychopathic traits to suicide-related behavior among offenders. *Law and Human Behavior*, *32*, 511–525. <http://dx.doi.org/10.1007/s10979-007-9122-8>
- Edens, J. F., Buffington, J. K., & Tomicic, T. L. (2000). An investigation of the relationship between psychopathic traits and malingering on the psychopathic personality inventory. *Assessment*, *7*, 281–296. <http://dx.doi.org/10.1177/107319110000700307>
- Edens, J. F., Hart, S. D., Johnson, D. W., Johnson, J. K., & Olver, M. E. (2000). Use of the Personality Assessment Inventory to assess psychopathy in offender populations. *Psychological Assessment*, *12*, 132–139. <http://dx.doi.org/10.1037/1040-3590.12.2.132>
- Edens, J. F., Marcus, D. K., Lilienfeld, S. O., & Poythress, N. G., Jr. (2006). Psychopathic, not psychopath: Taxometric evidence for the dimensional structure of psychopathy. *Journal of Abnormal Psychology*, *115*, 131–144. <http://dx.doi.org/10.1037/0021-843X.115.1.131>
- Edens, J. F., & Ruiz, M. A. (2006). On the validity of validity scales: The importance of defensive responding in the prediction of institutional misconduct. *Psychological Assessment*, *18*, 220–224. <http://dx.doi.org/10.1037/1040-3590.18.2.220>
- Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, *48*, 384–392.
- First, M. B., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). *Structured Clinical Interview for DSM-IV Axis II Personality Disorders*, (SCID-II). Washington, DC: American Psychiatric Press.
- Fritz, M. S., & MacKinnon, D. P. (2008). A graphical representation of the mediated effect. *Behavior Research Methods*, *40*, 55–60. <http://dx.doi.org/10.3758/BRM.40.1.55>
- Gacono, C. B., Meloy, J. R., Sheppard, K., Speth, E., & Roske, A. (1995). A clinical investigation of malingering and psychopathy in hospitalized insanity acquittees. *Bulletin of the American Academy of Psychiatry and the Law*, *23*, 387–397.
- Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychology, Public Policy, and Law*, *2*, 293–323. <http://dx.doi.org/10.1037/1076-8971.2.2.293>
- Hare, R. D. (1985). Comparison of procedures for the assessment of psychopathy. *Journal of Consulting and Clinical Psychology*, *53*, 7–16. <http://dx.doi.org/10.1037/0022-006X.53.1.7>
- Hare, R. D. (1991/2003). *Manual for the Revised Psychopathy Checklist* (1st ed.). Toronto, Ontario, Canada: Multi-Health Systems.
- Hart, S. D., Hare, R. D., & Forth, A. E. (1996). Psychopathy as validation of the Revised Psychopathy Checklist. In J. Monahan & H. J. Steadman (Eds.), *Violence and mental disorder: Developments in risk assessment* (pp. 81–98). Chicago, IL: University of Chicago Press.
- Hopwood, C. J., Ansell, E. B., Pincus, A. L., Wright, A. G., Lukowitsky, M. R., & Roche, M. J. (2011). The circumplex structure of interpersonal sensitivities. *Journal of Personality*, *79*, 707–740. <http://dx.doi.org/10.1111/j.1467-6494.2011.00696.x>
- John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality*, *61*, 521–551. <http://dx.doi.org/10.1111/j.1467-6494.1993.tb00781.x>
- Jones, S., & Miller, J. D. (2012). Psychopathic traits and externalizing behaviors: A comparison of self- and informant reports in the statistical prediction of externalizing behaviors. *Psychological Assessment*, *24*, 255–260. <http://dx.doi.org/10.1037/a0025264>
- Kelsey, K. R., Rogers, R., & Robinson, E. V. (2014). Self-report measures of psychopathy: What is their role in forensic assessments? *Journal of Psychopathology and Behavioral Assessment*. Advance online publication. <http://dx.doi.org/10.1007/s10862-014-9475-5>
- Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology*, *68*, 151–158. <http://dx.doi.org/10.1037/0022-3514.68.1.151>
- Li, A., & Bagger, J. (2006). Using the BIDR to distinguish the effects of impression management and self-deception on the criterion validity of personality measures: A meta-analysis. *International Journal of Selection and Assessment*, *14*, 131–141. <http://dx.doi.org/10.1111/j.1468-2389.2006.00339.x>

- Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal populations. *Journal of Personality Assessment*, *66*, 488–524. [http://dx.doi.org/10.1207/s15327752jpa6603\\_3](http://dx.doi.org/10.1207/s15327752jpa6603_3)
- Lilienfeld, S. O., & Fowler, K. A. (2006). The self-report assessment of psychopathy. In C. Patrick (Ed.), *Handbook of psychopathy* (pp. 107–132). New York, NY: Guilford Press.
- Lilienfeld, S. O., Patrick, C. J., Benning, S. D., Berg, J., Sellbom, M., & Edens, J. F. (2012). The role of fearless dominance in psychopathy: Confusions, controversies, and clarifications. *Personality Disorders: Theory, Research, and Treatment*, *3*, 327–340. <http://dx.doi.org/10.1037/a0026987>
- Lykken, D. T. (1957). A study of anxiety in the sociopathic personality. *Journal of Abnormal and Social Psychology*, *55*, 6–10. <http://dx.doi.org/10.1037/h0047232>
- Marcus, D. K., Fulton, J. J., & Edens, J. F. (2013). The two-factor model of psychopathic personality: Evidence from the psychopathic personality inventory. *Personality Disorders: Theory, Research, and Treatment*, *4*, 67–76. <http://dx.doi.org/10.1037/a0025282>
- McCord, W., & McCord, J. (1964). *The psychopath: An essay on the criminal mind*. Oxford, England: D. Van Nostrand.
- McCrae, R. R., & Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology*, *51*, 882–888. <http://dx.doi.org/10.1037/0022-006X.51.6.882>
- McCrae, R. R., & Costa, P. T., Jr. (1989). The structure of interpersonal traits: Wiggins's circumplex and the five-factor model. *Journal of Personality and Social Psychology*, *56*, 586–595. <http://dx.doi.org/10.1037/0022-3514.56.4.586>
- McGrath, R. E., Mitchell, M., Kim, B. H., & Hough, L. (2010). Evidence for response bias as a source of error variance in applied assessment. *Psychological Bulletin*, *136*, 450–470. <http://dx.doi.org/10.1037/a0019216>
- McHoskey, J. W., Worzel, W., & Szyarto, C. (1998). Machiavellianism and psychopathy. *Journal of Personality and Social Psychology*, *74*, 192–210. <http://dx.doi.org/10.1037/0022-3514.74.1.192>
- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis, MN: University of Minnesota Press. <http://dx.doi.org/10.1037/11281-000>
- Meehl, P. E. (1957). When shall we use our heads instead of the formula? *Journal of Counseling Psychology*, *4*, 268–273. <http://dx.doi.org/10.1037/h0047554>
- Miller, J. D., Jones, S. E., & Lynam, D. R. (2011). Psychopathic traits from the perspective of self and informant reports: Is there evidence for a lack of insight? *Journal of Abnormal Psychology*, *120*, 758–764. <http://dx.doi.org/10.1037/a0022477>
- Miller, J. D., & Lynam, D. R. (2012). An examination of the Psychopathic Personality Inventory's nomological network: A meta-analytic review. *Personality Disorders: Theory, Research, and Treatment*, *3*, 305–326. <http://dx.doi.org/10.1037/a0024567>
- Morey, L. C. (1991). *Personality Assessment Inventory (PAI)*. Odessa, FL: PAR.
- Morey, L. C. (2012). Detection of response bias in applied assessment: Comment on McGrath et al. (2010). *Psychological Injury and Law*, *5*, 1–9.
- Newman, J. P., & Kosson, D. S. (1986). Passive avoidance learning in psychopathic and nonpsychopathic offenders. *Journal of Abnormal Psychology*, *95*, 252–256. <http://dx.doi.org/10.1037/0021-843X.95.3.252>
- Newman, J. P., Widom, C. S., & Nathan, S. (1985). Passive avoidance in syndromes of disinhibition: Psychopathy and extraversion. *Journal of Personality and Social Psychology*, *48*, 1316–1327. <http://dx.doi.org/10.1037/0022-3514.48.5.1316>
- Nielsen, I. J., Nentjes, L., Merckelbach, H., & Bernstein, D. P. (2015). Antisocial features and “faking bad”: A critical note. *International Journal of Law and Psychiatry*. Advanced online publication. <http://dx.doi.org/10.1016/j.ijlp.2015.03.005>
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red her-
- ring. *Journal of Applied Psychology*, *81*, 660–679. <http://dx.doi.org/10.1037/0021-9010.81.6.660>
- Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, *66*, 1025–1060. <http://dx.doi.org/10.1111/1467-6494.00041>
- Poythress, N. G., Lilienfeld, S. O., Skeem, J. L., Douglas, K. S., Edens, J. F., Epstein, M., & Patrick, C. J. (2010). Using the PCL-R to help estimate the validity of two self-report measures of psychopathy with offenders. *Assessment*, *17*, 206–219. <http://dx.doi.org/10.1177/1073191109351715>
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*, 717–731. <http://dx.doi.org/10.3758/BF03206553>
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879–891. <http://dx.doi.org/10.3758/BRM.40.3.879>
- Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods*, *16*, 93–115. <http://dx.doi.org/10.1037/a0022658>
- Ray, J. V., Hall, J., Rivera-Hudson, N., Poythress, N. G., Lilienfeld, S. O., & Morano, M. (2013). The relation between self-reported psychopathic traits and distorted response styles: A meta-analytic review. *Personality Disorders: Theory, Research, and Treatment*, *4*, 1–14. <http://dx.doi.org/10.1037/a0026482>
- Rogers, R., & Cruise, K. R. (2000). Malingering and deception among psychopaths. In C. B. Gacono (Ed.), *The clinical and forensic assessment of psychopathy: A practitioner's guide*. The LEA series in personality and clinical psychology (pp. 269–284). Mahwah, NJ: Erlbaum.
- Sackeim, H. A., & Gur, R. C. (1979). Self-deception, other-deception, and self-reported psychopathology. *Journal of Consulting and Clinical Psychology*, *47*, 213–215. <http://dx.doi.org/10.1037/0022-006X.47.1.213>
- Sellbom, M. (2011). Elaborating on the construct validity of the Levenson self-report psychopathy scale in incarcerated and non-incarcerated samples. *Law and Human Behavior*, *35*, 440–451. <http://dx.doi.org/10.1007/s10979-010-9249-x>
- Smith, S. F., Watts, A. L., & Lilienfeld, S. O. (2014). On the trail of the elusive successful psychopath. *The Psychologist*, *27*, 506–511.
- Tellegen, A. (1982). *Brief manual for the Multidimensional Personality Questionnaire*. Unpublished manuscript, University of Minnesota, Minneapolis.
- Uziel, L. (2010). Rethinking social Desirability scales from impression management to interpersonally oriented self-control. *Perspectives on Psychological Science*, *5*, 243–262. <http://dx.doi.org/10.1177/1745691610369465>
- Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality*, *40*, 472–481. <http://dx.doi.org/10.1016/j.jrp.2005.03.003>
- Verona, E., Patrick, C. J., & Joiner, T. E. (2001). Psychopathy, antisocial personality, and suicide risk. *Journal of Abnormal Psychology*, *110*, 462–470.
- Verschuere, B., Uzieblo, K., De Schryver, M., Douma, H., Onraedt, T. Y., & Crombez, G. (2014). The inverse relation between psychopathy and faking good: Not response bias but true variance in psychopathic personality. *Journal of Forensic Psychiatry & Psychology*, *25*, 705–713. <http://dx.doi.org/10.1080/14789949.2014.952767>
- Widom, C. S. (1977). A methodology for studying noninstitutionalized psychopaths. *Journal of Consulting and Clinical Psychology*, *45*, 674–683.

Received November 12, 2014

Revision received April 20, 2015

Accepted April 21, 2015 ■