Personality Disorders: Theory, Research, and Treatment
2013, Vol. 4, No. 1, 1–14

1949-2715/12/$12.00 DOI: 10.1037/a0026482

The Relation Between Self-Reported Psychopathic Traits and Distorted Response Styles: A Meta-Analytic Review

James V. Ray and Jason Hall
University of South Florida

Nicole Rivera-Hudson
University of Southern Mississippi

Norman G. Poythress
University of South Florida

Scott O. Lilienfeld
Emory University

Mario Morano
University of South Florida

A concern among researchers is that self-report measures may not be valid indicators of psychopathic traits due to the core features of psychopathy (e.g., lying, deception/manipulation). The current study addresses this issue by combining effects sizes from studies published on or before March 31, 2010 to examine the relation between scores of 3 widely used self-report psychopathy measures—the Psychopathic Personality Inventory (PPI; Lilienfeld & Andrews, 1996) and its revised version (PPI-R; Lilienfeld & Widows, 2005) and Levenson's Self-Report Psychopathy scale (LSRP; Levenson, Kiehl, & Fitzpatrick, 1995) and scores on measures assessing response style (i.e., faking good and faking bad). Effect sizes were obtained from 45 studies for total, Factor 1, and Factor 2 scores (faking good: $k = 54$, 55, and 55, respectively; faking bad: $k = 51$, 50, and 50, respectively). Based on a random effects model, a significant negative association was found between social desirability/faking good and both total ($r_w = -.11$, $p < .01$) and F2 ($r_w = -.16$, $p < .01$) scores, and moderation analyses suggested that effect sizes varied as a function of psychopathy scale and validity scale used. Significant positive associations were also found between faking bad and both total ($r_w = .27$, $p < .05$) and F2 ($r_w = .32$, $p < .05$) scores. Also, moderation analyses suggested that effect sizes varied as a function of study location, psychopathy scale, and validity scale. Despite several limitations (e.g., inclusion of only published studies, limited moderators, exclusion of other measures), the general findings temper concerns of positive response bias and underscore the validity of self-report psychopathy scales.

Keywords: self-report, psychopathy, response style, Psychopathic Personality Inventory, Levenson's Self-Report Psychopathy scale

Wanton and effective use of deception is regarded as a hallmark feature of psychopathy, a personality disorder that entails a constellation of distinctive affective-interpersonal features (e.g., superficial charm, deception/manipulation, lack of empathy/remorse) in the context of a chronically disorganized and often antisocial lifestyle. A related assumption is that psychopathic individuals—particularly those exhibiting prominent affective-interpersonal features of the disorder—are more likely than others to either (a) present themselves in a positive light or to (b) malinger psychiatric symptoms in the context of clinical or personality assessment (Book, Holden, Starzyk, Wasylkiw, & Edwards, 2006; Edens, Buffington, & Tomicic, 2000; MacNeil & Holden, 2006; see Lilienfeld, 1994; Lilienfeld & Fowler, 2006 for a discussion).

Indeed, the assumption that the self-report detection of psychopathy is untrustworthy has become a virtual truism to some and has almost surely contributed to widespread skepticism concerning the use of self-report measures in the detection of psychopathy. For example, Hart, Hare, and Forth (1994) argued 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). Moreover, Edens, Hart, et al. (2000), argued that ". . . self-reports may be particularly susceptible to response distortion. This is a major potential problem, because deceitfulness is construed as a core symptom of psychopathy" (p. 137). In contrast, a recent study revealed considerable correspondence between self- and informant-reports of psychopathy, suggesting that individuals with psychopathic traits are often both willing and capable of providing accurate assessments of themselves on such traits (Miller, Jones, & Lynam, 2011).

Deceptive response styles typically fall into two general categories: malingering and socially desirable responding. Malingering refers to attempts to feign psychiatric symptoms or mental illness. Social desirability, however, is more heterogeneous and includes

This article was published Online First January 23, 2012.

James V. Ray, Jason Hall, and Norman G. Poythress, Department of Mental Health Law & Policy, Florida Mental Health Institute, University of South Florida; Nicole Rivera-Hudson, Department of Psychology, University of Southern Mississippi; Scott O. Lilienfeld, Department of Psychology, Emory University; Mario Morano, Department of Psychology, University of South Florida.

James V. Ray is now with the Department of Psychology at University of New Orleans.

Correspondence concerning this article should be addressed to James V. Ray, Department of Psychology, University of New Orleans, New Orleans, LA 70148. E-mail: jvray@uno.edu

positive impression management (i.e., intentionally and knowingly presenting oneself in a positive light) and self-deception (i.e., internalization of positive reports of one's self; see, e.g., Paulhus, 1984). The former is typically assumed to be conscious and obvious, the latter largely unconscious and subtle.

A number of instruments have been developed and validated to detect impression management, malingering, or other aberrant response styles (e.g., Balanced Inventory of Deceptive Responding [BIDR; Paulhus, 1984]; Marlowe-Crowne Social Desirability Scale [Crowne & Marlowe, 1960]; The Structured Interview of Reported Symptoms [SIRS; Rogers, Bagby, & Dickens, 1992]), and are widely used in various contexts. Nevertheless, the research literature on the relationship between psychopathy and response styles, including impression management and malingering, is relatively sparse. Furthermore, recent years have witnessed the emergence of several widely used self-report psychopathy instruments (see Lilienfeld & Fowler, 2006, for a review). Despite promising support for the validity of such measures, one serious and lingering concern is that psychopathic individuals are more likely than others to adopt distorted response styles on self-report inventories. This concern raises questions regarding the increasingly popular practice of detecting psychopathic traits using such indices.

In light of these issues, the aim of the present study was to address—via a meta-analytic review of published findings—two related questions: (a) Are self-reported psychopathic traits systematically associated with tendencies to dissimulate? (b) Do the two major factors of psychopathy measures (see below) relate differentially to distorted response styles?

Facets of Psychopathy and Distorted Response Styles

The Psychopathy Checklist-Revised (PCL-R; Hare, 1991/2003) is the predominant instrument for assessing psychopathic features, and a wealth of data supports its reliability and validity (Hare, 1991/2003). Structural analysis of the PCL-R items has consistently revealed a multifactorial structure, with the most common model being one in which the PCL-R is underpinned by two moderately correlated but distinct factors: an affective-interpersonal factor (Factor 1), comprising such features as superficial charm, narcissism, use of deception/manipulation, lack of empathy/remorse, blame externalization, and shallow affective experience; and an antisocial deviance factor (Factor 2), consisting of items tapping impulsivity, irresponsibility, boredom proneness, intentional dependence on others, lack of a coherent life plan, aggression, and early onset, persistent, and varied criminal activity (but see Hare, 1991/2003, and Cooke & Michie, 2001, for competing four and three factor models, respectively).

Psychopaths are commonly believed to be more prone to exaggerate or fabricate psychiatric symptoms when it is strategically useful (as is often the case in forensic assessment contexts, such as competency or insanity proceedings) compared with when there is no incentive to do so (Rogers & Cruise, 2000). Many also assume that individuals with high scores on psychopathy measures are more likely than those with low scores to engage in this response style (Gacono, Meloy, Sheppard, Speth, & Roske, 1995; alternatively, see Poythress, Edens, & Watkins, 2001). In contrast, others argue that under certain circumstances (i.e., in the context of a job application), psychopaths are more likely than nonpsychopaths to engage in positive impression management (Book et al., 2006). Psychopaths are also believed by many to engage in deception even when there is no clear motivation to do so (see Ekman, 1993). Given the insincere, deceptive, and manipulative interpersonal style embodied by PCL-R Factor 1, it is widely held that this factor in particular entails tendencies to adopt a deceptive stance toward assessment measures across various contexts (Book et al., 2006; Rogers & Cruise, 2000). In spite of these assumptions, little research has focused on relations between the facets of psychopathy and distorted response styles.

Research using the PCL-R has suggested that psychopathy is associated with an increased likelihood of malingering psychiatric symptoms, although findings regarding the two psychopathy factors have been inconclusive. For example, Gacono et al. (1995) reported higher scores on both PCL-R factors among hospitalized insanity plea acquittees known to have malingered psychiatric symptoms compared with those who did not. More recently, Kucharski, Duncan, Egan, and Falkenbach (2006) reported that psychopaths undergoing court-ordered forensic assessment produced higher scores on several measures of malingering; additionally, scores on PCL-R Factor 1 (but not Factor 2) significantly predicted malingering classification, albeit with poor sensitivity and specificity. It thus remains unclear to what extent psychopathy factors relate to tendencies toward distorted patterns of responding to assessment measures.

Self-Report Psychopathy and Response Styles

Although it is by far the most widely used and best-validated psychopathy measure, the PCL-R entails a lengthy (typically 90 min) clinical interview by a highly trained rater, supplemented by a detailed review of collateral file information. The PCL-R was designed specifically for assessing psychopathic traits in offender populations. Thus, the PCL-R is often not practical or well-suited for efficient large-scale screening or research involving nonforensic samples.

In response to these limitations, several self-report measures of psychopathic traits have emerged in recent years. The most widely used measures are the Psychopathic Personality Inventory (PPI; Lilienfeld & Andrews, 1996), its revision (PPI-R; Lilienfeld & Widows, 2005), the Self-Report Psychopathy Scale (SRP-II; Hare, 1990), and the Levenson Self-Report Psychopathy Scale (LSRP; Levenson, Kiehl, & Fitzpatrick, 1995). Accumulating evidence supports the validity of the self-report approach to the assessment of psychopathy in community and forensic settings (for a review, see Lilienfeld & Fowler, 2006). For instance, both the PPI and LSRP often exhibit a two-factor structure broadly similar to the PCL-R (Benning, Patrick, Hicks, Blonigen, & Krueger, 2003; Levenson et al., 1995; although see Neumann, Malterer, & Newman, 2008), show at least moderate convergent validity with the PCL-R (Berardino, Meloy, Sherman, & Jacobs, 2005; Brinkley, Schmidt, Smith, & Newman, 2001; Edens, Poythress, Lilienfeld, & Patrick, 2008; Malterer, Lilienfeld, Neumann, & Newman, 2010), and exhibit patterns of correlations with external criteria that are consistent with theoretical prediction (Benning, Patrick, Salekin, & Leistico, 2005; Lynam, Whiteside, & Jones,

1999; Miller, Gaughan & Pryor, 2008; Patrick, Edens, Poythress, Lilienfeld, & Benning, 2006; Poythress et al., 2010).

One concern that has yet to be resolved is the issue of deceptive responding in self-report psychopathy assessment. Although conceptual issues regarding the problem of deception in self-report psychopathy assessment have been reviewed elsewhere (Lilienfeld, 1994; Lilienfeld & Fowler, 2006), few studies have addressed this issue empirically. Initial studies have found, contrary to widespread expectation (e.g., Rogers & Cruise, 2000), that scores on self-report psychopathy measures typically demonstrate modest *negative* correlations with tendencies toward positive impression management or socially desirable response styles (Lilienfeld & Andrews, 1996) and are not significantly related to success at malingering psychiatric symptoms (Edens, Buffington, et al., 2000; Poythress et al., 2001). Edens, Buffington, et al. (2000) did find, however, that individuals with elevated self-reported psychopathy scores reported both greater willingness to malinger and confidence in their abilities to do so successfully. Further, although the PPI and PPI-R include validity scales designed to detect distorted response styles, these scales are most often used to identify grossly invalid profiles on the basis of prescribed cutoff scores. Consequently, correlations between validity and trait scales (reflecting more subtle but systematic covariation between response styles and psychopathic traits) are infrequently the focus of study, and rarely reported.

Nevertheless, numerous studies that administered self-report psychopathy scales collected data on response styles, including deviant responding or malingering, and positive impression management or socially desirable responding. In the case of studies that used the PPI/PPI-R (henceforth referred to jointly as the "PPI/R"), which include internal validity scales, correlations between response style indicators and psychopathic traits are straightforward to obtain. Although the LSRP does not include built-in validity indices, several studies have used this measure in conjunction with either stand-alone response style measures, or with omnibus self-report psychological inventories that include scales designed to detect deviant or socially desirable response styles.

The Present Study

In light of the issues we have reviewed, the aims of the present study were to (a) examine the degree to which self-reported psychopathic traits relate to indices of response styles, namely, socially desirable responding and malingering and (b) evaluate differential relationships between the two psychopathy factors and response styles. We conducted a meta-analytic review of published studies that administered both a self-report psychopathy measure (either the PPI/R or LSRP) and one or more measures designed to detect response styles, especially socially desirable responding and positive impression management (i.e., "faking good") and/or malingering (i.e., "faking bad").[1] In addition, we tested for the potential moderating effects of several factors, including psychopathy instrument (i.e., PPI/R vs. LSRP), response validity indices used, sample characteristics (i.e., gender, community vs. offender), and study location (i.e., U.S. vs. Canada vs. Europe).

Method

Measures

Psychopathy scales.

Psychopathic Personality Inventory (Lilienfeld & Andrews, 1996). The PPI is a 187-item self-report measure of psychopathic features. Individual responses are based on a 4-point Likert-type scale (1 = False – 4 = True). The PPI consists of 8 content scales, 7 of which load on two higher order factors: PPI-I (labeled Fearless Dominance [FD]) and PPI-II (labeled Impulsive Antisociality [IA]; Benning et al., 2003). The eighth content scale, Coldheartedness, did not load highly on either factor. Excellent internal consistency has also been reported for the PPI Total and content scales in both student (Lilienfeld & Andrews, 1996) and offender (Poythress et al., 2010) samples.

Psychopathic Personality Inventory – Revised (Lilienfeld & Widows, 2005). The PPI-R is a 154-item revised version of the PPI. The PPI-R was developed to address concerns regarding administration time, reading level, and potential cultural bias in the original PPI, yet follows the same format (i.e., 4-point Likert-type scale). The PPI-R retains the same structural properties as the PPI, with 8 content scales that yield a total psychopathy score; seven of the 8 content scales load on two higher order factors, FD and Self-Centered Impulsivity (SCI; Lilienfeld & Widows, 2005). The Coldheartedness content scale does not load on either factor. The internal consistency of the PPI-R among community/college samples is adequate (α = .78−.92) and the PPI-R has demonstrated promising convergent validity with other self-report measures of psychopathy (Lilienfeld & Widows, 2005).

Levenson's Self-Report Psychopathy Scale (Levenson et al., 1995). The LSRP is a 26-item self-report measure of psychopathic personality traits designed for use in noninstitutionalized settings. Responses are based on a 4-point Likert-type scale (1 = agree strongly – 4 = disagree strongly). The LSRP yields a Total score and two factor scores, labeled Primary and Secondary psychopathy. The LSRP has shown good internal consistency and construct validity (Jakobwitz & Egan, 2006; Lynam et al., 1999).

Validity scales.

Balanced Inventory of Desirable Responding (Paulhus, 1984). The BIDR is a 40-item self-report instrument that measures deceptive responding. Responses are based on a 7-point Likert-type scale (1 = not true – 7 = very true) with higher scores indicating deceptive responding. All 40 items can be summed to reflect an overall score of socially desirable responding. Additionally, 20 items can be summed to reflect an impression management scale (conscious self-presentation to an audience), and the remaining 20 items can be summed to reflect a positive self-deception scale. In this study we limited analyses to the overall social desirability index.

The Marlowe-Crowne Social Desirability Scale (MCSDS; Crowne & Marlowe, 1960). The MCSDS is a 33-item true or false self-report instrument used to assess positive impression

[1] The decision to focus the current analyses on the two versions of the PPI and the LSRP was based on the volume of studies that have employed these measures. The SRP was not considered in the current study due to its multiple versions that have evinced inconsistent factor structures (see e.g., Williams & Paulhus, 2004).

management. Eighteen of the items ask respondents about socially acceptable yet uncommon behaviors; endorsement of each item is indicative of socially desirable responding. The remaining 15 items ask about socially unacceptable yet common practices; endorsing each item on this scale is indicative of defensive responding.

Personality Assessment Inventory (PAI) - Negative Impression Management (NIM) and Positive Impression Management (PIM) scales (Morey, 1991). All items on the PAI are based on a 4-point Likert-type scale (0 = false – 3 = very true). The 9 items that compose the PAI-NIM present symptoms that are out of the ordinary or rarely experienced to assess respondents' tendencies to present themselves in a negative light. Higher scores on the PAI-NIM ostensibly indicate a tendency to "fake bad." The PAI-PIM scale includes 9 items that assess the likelihood that one will respond defensively and present oneself in a positive light. Higher scores on the PAI-PIM ostensibly indicate a tendency to "fake good."

PPI Deviant Responding (DR) and Unlikely Virtues (UV) scales (Lilienfeld & Andrews, 1996). The DR and UV scales are embedded within the PPI (see earlier description). The UV scale (14 items) is identical to the UV scale of the Multidimensional Personality Questionnaire (MPQ; Tellegen, 1978/1982) and was incorporated into the PPI to detect individuals who have a tendency engage in positive impression management. The DR scale (10 items) consists of bizarre items that are unlikely to be observed in genuine psychopathology and was designed to detect individuals who are faking bad, responding inconsistently or carelessly, or who have trouble understanding item content.

PPI-R DR and Virtuous Responding (VR) scales (Lilienfeld & Widows, 2005). The DR and VR scales are embedded within the PPI-R. The DR ("faking bad"; 10 items) and VR ("faking good"; 13 items) scales of the PPI-R are also akin to the validity scales included in the original PPI (DR and UV, respectively).

The Psychological Inventory of Criminal Thinking Styles (PICTS) Confusion (Cf) and Defensiveness (Df) scales (Walters, 1995). All items on the PICTS are based on a 4-point Likert-type scale (1 = disagree – 4 = strongly agree). Each validity scale is composed of 8 items. High scores on the Cf scale are indicative of "faking bad" or malingering, whereas high scores on the Df scale are indicative of "faking good" or responding in a socially desirable manner.

Study Selection

A literature search was conducted to locate all published studies through March 31, 2010 that employed at least one of three measures (LSRP, PPI, and PPI-R). Two strategies were used to locate pertinent articles. First, four article databases were thoroughly searched: PubMed, PsycInfo, Science Direct, and Medline. Each database was searched for articles using various combinations of relevant search terms (e.g., LSRP, Levenson, PPI, self-report psychopathy scales). Also, each database was searched for all articles that cited the original source of each measure. Second, the reference section of each article retrieved was searched manually for any potentially relevant studies.

Once retrieved, each article was screened to determine its appropriateness for inclusion. An article was marked for inclusion in the meta-analysis if (a) at least one of the three psychopathy measures was administered and (b) at least one relevant validity scale was administered. Articles meeting both inclusion criteria were examined to determine if correlations between validity scales and psychopathy scores (i.e., total and factor scores) were reported. If these correlations were not reported, but the relevant psychopathy and validity scales had been administered, the corresponding author was contacted and asked to provide this information. Thus, only correlations from published studies were used in the current review, although many of the correlations themselves were unpublished.

We initially retrieved a total of 175 articles; 92 of these studies met the aforementioned inclusion criteria. Of these articles, 10 were excluded from the meta-analysis because data were reported for samples that were used in more than one publication, yielding a total of 82 articles. Of those 82 articles, only three reported correlations between psychopathy measures and validity indices; the corresponding authors for the remaining 79 articles were contacted to request information. In response to our initial request and a 6-week follow-up, we received enough responses to yield a final sample size of 45 studies.

Coding Strategy

The information collected from each of the 45 studies included the following: (a) the self-report psychopathy measure administered; (b) the validity scale administered; (c) gender characteristics of the sample; (d) racial composition of sample; (e) the broader population from which the sample was drawn (e.g., undergraduate, prisoner, forensic); and (f) correlations between validity and psychopathy measures (including factor scores, if available). In all, 23 studies used the PPI, 12 used the PPI-R, and 14 used the LSRP. However, several studies contributed multiple effect sizes because multiple psychopathy measures were administered. There were a total of 127 effect sizes for total psychopathy, and 122 for each of the psychopathy factor scores. When studies reported data on multiple validity scales of the same type (i.e., social desirability/faking good or faking bad), only correlations for the most commonly used validity scale (within this sample of studies) were included in the analyses to avoid violating assumptions of independent effect sizes. This approach was taken (as opposed to averaging the effect sizes) to preserve maximum possible homogeneity among criterion measures. This approach resulted in 105 effect sizes for total psychopathy (54 for social desirability/faking good, 51 for faking bad), 105 effect sizes for PPI-I/LSRP-Primary (55 for social desirability/faking good, 50 for faking bad), and 105 for PPI-II/LSRP-Secondary (55 for social desirability/faking good, 50 for faking bad) included in the final analyses.

Analytic Approach

Separate analyses were conducted for validity measures assessing social desirability/faking good and faking bad response styles. Additionally, mean effect sizes reflecting correlations between psychopathy and validity measures were calculated separately for total and factor scores (F1 = PPI/R-I and LSRP-Primary; F2 = PPI/R-II and LSRP-Secondary). For each of these effect sizes, the Q and I^2 statistics were examined to determine if there was significant variation among the effect sizes beyond what could be attributed to sampling error. In the cases in which significant

heterogeneity among effect sizes was indicated, several potential moderating variables were examined:

Gender. Sample gender characteristics associated with each effect size were recorded, and coded as either male, female, or mixed. In some cases, effect sizes were reported separately for males and females and were recorded as such. Those articles that reported on data for mixed samples (and did not provide separate correlations for males and females) were not included in the moderation analyses for gender.

Sample type. Study population was initially coded as prisoner ($n = 49$), forensic ($n = 6$), clinical ($n = 1$), community ($n = 21$), undergraduate ($n = 45$), or mixed ($n = 3$); however, due to a low number of studies in some cells, these categories were collapsed into *community* (community, undergraduate, or mixed) and *institutional* (prisoner, forensic, or clinical) categories.

Study location. Because of potential cross-cultural differences in the correlates of psychopathy (see, e.g., Sullivan & Kosson, 2006), study location was included as a moderator. Geographic location was coded based on where samples were drawn: United States ($n = 86$), Canada ($n = 7$), Europe ($n = 30$), or other ($n = 2$). Given the relatively small number of studies conducted outside the U.S., this variable was later collapsed into *U.S.* and *non-U.S.* categories for moderation analyses.

Psychopathy scale. The self-report measure used to operationalize psychopathy and its facets was examined as a potential moderator. Psychopathy measure was coded as PPI, PPI-R, or LSRP.[2]

Validity scale. The validity scale(s) administered in each study was coded for moderation analyses. Ten different "social desirability/faking good" scales and five different "faking bad" scales were employed in the studies. However, as mentioned above, several studies used multiple validity scales, and in such cases only the most commonly used measure was retained in the moderator analyses. Additionally, two of the scales assessing social desirability/faking good contributed only one effect size for Total, F1, or F2. Due to computational limitations, those studies with only one effect size could not be included in the moderator analysis, yielding a total of five social desirability/faking good scales and four faking bad scales.

Statistical Procedures and Software

All analyses were conducted using SPSS version 19. First, mean effect sizes were calculated by transforming all Pearson r's to Fisher's z scores (z_r) and computing their corresponding inverse variance weights (i.e., $n - 3$). The resulting z_r values were aggregated using the meta-analysis macro designed for use with SPSS (*MeanES*; Lipsey & Wilson, 2001; Wilson, 2005). The macro follows guidelines outlined by Hedges and Olkin (1985) and calculates mean effect sizes, as well as upper and lower 95% confidence intervals, using both fixed and random effects based on the inverse variance weight of each effect size. The macro also conducts a test of homogeneity using the Q statistic (Hedges & Olkin, 1985). The Q statistic is distributed as a chi-square with $k - 1$ degrees of freedom, where k equals the number of effect sizes. Significant Q values suggest heterogeneity among the effects sizes beyond what would be expected due to sampling error. Because the power of the Q statistic may be attenuated when the number of effect sizes is small, the I^2 statistic was also calculated ($I^2 = [Q - df]/Q$). The I^2 reflects the proportion of variance that can be attributed to heterogeneity among studies (Higgins & Thompson, 2002). Higgins and Thompson (2002) suggested that I^2 values greater than .50 indicate noteworthy heterogeneity.

Moderation analyses were conducted to follow up on significant heterogeneity tests. Specifically, because all moderators were categorical in nature, the analog to analysis of variance (ANOVA) was conducted using the *MetaF* macro (Lipsey & Wilson, 2001; Wilson, 2005). The analog-to-ANOVA partitions the variability represented in the Q statistic into two components: within (Q_W) and between (Q_B). A significant Q_B indicates that the means across groups of effect sizes differ significantly, suggesting that the grouping variable (i.e., moderator) accounts for a significant amount of the heterogeneity in the effect sizes. The analog-to-ANOVA also reports a Q_w for each group. A significant Q_w suggests that the effect sizes within that subgroup are not homogeneous.

To account for potential publication bias, Orwin's (1983) fail-safe N was calculated for each significant mean effect size in both the overall mean effect size analyses and moderation analyses. Orwin's fail-safe N indicates the number of studies with null findings (i.e., nonsignificant correlation coefficients) that would be needed to reduce a significant effect size to a given size, in this case, to the size needed to render the finding nonsignificant.

Results

Social Desirability/Faking Good

Study characteristics and correlations reported between psychopathy trait scores and faking good for each study are reported in Appendix 1. Initial inspection of the data suggested the possible presence of outlier effect sizes; thus, all analyses we report were also conducted after removing outliers that were 2 SDs above the mean as suggested by Lipsey and Wilson (2001); effect sizes considered outliers for social desirability/faking good are identified in Appendix 1. No considerable changes in the I^2 statistic were observed in analyses with and without outliers, and the following analyses were conducted including outliers to retain as many effect sizes as possible.

Mean Effect Sizes

Table 1 presents mean effect sizes for correlations between psychopathy and social desirability/faking good measures for Total, F1 (i.e., PPI/R-I and LSRP-Primary), and F2 (i.e., PPI/R-II and LSRP-Secondary) psychopathy trait scores based on random effects models, which assume that at least some of the heterogeneity in effect sizes across studies are due to methodological differences. A significant negative relationship was found between Total and F2 psychopathy scores and social desirability/faking good; overall, F1 was not significantly related to social desirability/faking good. Also, the large fail-safe N's reported in Table 1 for total and F2 scores suggest that a large number of unpublished studies with

[2] Because of the overlap between the PPI and the PPI-R (see Ray, Weir, Poythress, & Rickelm, 2011) these two measures were collapsed into one category (i.e., PPI/R).

Table 1
Mean Effect Sizes for Social Desirability/Faking Good

Psychopathy Scale	k (k_0)	r_u	95% CI_u	r_w	95% CI_w	Q (I^2)
Total Psychopathy	54 (243)	−.11	[−.36, .15]	−.11**	[−.18, −.04]	872.16 (.94)***
PPI-I/Primary	55 (–)	.01	[−.25, .27]	−.00	[−.06, .05]	494.31 (.89)***
PPI-II/Secondary	55 (385)	−.16	[−.40, .10]	−.16**	[−.25, −.06]	1703.65 (.97)***

Note. k = number of effect sizes; k_0 = Orwin's fail-safe N; r_u = unweighted mean effect size; CI_u = confidence interval for unweighted mean effect size; r_w = weighted mean effect size; CI_w = confidence interval for weighted mean effect size; Q/ I^2 = statistics for test of homogeneity (significant test indicates rejection of hypothesis of homogeneity).
** $p < .01$. *** $p < .001$.

nonsignificant findings (k_0 = 243 and 385) would be needed to reduce the mean effect sizes to nonsignificance. Both the Q and I^2 statistics indicated significant heterogeneity among effect sizes used to calculate the mean effect sizes for Total, F1 and F2 correlations with social desirability/faking good indices; thus, moderation analyses were performed to interpret the sources of this heterogeneity.

Moderation Analyses

Results of moderation analyses (random effects models) for social desirability/faking good are presented in Table 2. Analog-to-ANOVA tests revealed that the psychopathy scale was a significant moderator of the association between psychopathy and social desirability/faking good, but for F1 scores only (Q_b = 9.05, $p < .01$), indicating that the negative association between LSRP Primary scores and social desirability/faking good (r_w = −.11, $p < .05$) was significantly stronger than for PPI/R-I (r_w = .05, $p = .13$). All of the associations between LSRP scores and social desirability/faking good indices were significant and negative, whereas for PPI/R scores, only F2 exhibited a significant negative relationship with social desirability/faking good (r_w = −.11, $p < .05$). As shown in Table 2, tests of homogeneity revealed no evidence of significant variability among effect sizes within any of the Psychopathy scale × Psychopathy factor cells.

Validity scale was a significant moderator of the association between psychopathic traits and social desirability/faking good, for Total (Q_B = 20.47, $p < .001$), F1 (Q_B = 15.28, $p < .01$), and F2 (Q_B = 9.67, $p < .05$) scores. Whereas the PICTS-Df scale was negatively related to all psychopathy scores (Total, F1, and F2), the PPI-UV scale related only to Total and F2 scores; PAI-PIM was associated negatively with F1, and neither the PPI-R-VR nor the MCSDS were related to any psychopathy trait measures. Because the number of effect sizes contributing to analyses of MCSDS (k = 2), PAI-PIM (k = 3), and PICTS-Df (k = 4) were relatively small, findings for these measures should be interpreted cautiously. Tests of homogeneity conducted for each subgroup of effect sizes (comprising the 12 Validity scale × Psychopathy factor cells) suggested homogeneity within each cell, with the exception of the PPI-R-VR × Total psychopathy cell. However, subsequent moderation analyses on effect sizes within this cell did not reveal any significant moderators of the association between PPI-R-VR and Total psychopathy scores. No significant moderating effects on the association between psychopathy trait scales and social desirability/faking good indices were found for sample

Table 2
Moderation Analysis for Social Desirability/Faking Good

	Total				PPI-I/Primary				PPI-II/Secondary			
Moderator	k (k_0)	r_w	95% CI	Q_w (I^2)	k (k_0)	r_w	95% CI	Q_w (I^2)	k	r_w	95% CI	Q_w (I^2)
Psych Scale												
PPI/R	38 (–)	−.08	[−.16, .00]	33.78 (.00)	38 (–)	.05	[−.01, .11]	37.93 (.02)	38 (171)	−.11*	[−.22, −.00]	40.88 (.09)
LSRP	16 (80)	−.18**	[−.30, −.06]	19.39 (.17)	17 (45)	−.11*	[−.19, −.02]	19.06 (.16)	17 (125)	−.25**	[−.39, −.09]	14.48 (.00)
Validity Scale												
PAI-PIM	3 (–)	−.23	[−.45, .01]	2.39 (.16)	3 (5)	−.21*	[−.41, −.02]	4.25 (.53)	3 (–)	−.25	[−.55, .11]	3.84 (.48)
PPI-UV	25 (200)	−.18***	[−.26, −.09]	10.35 (.00)	24 (–)	.03	[−.04, .11]	22.38 (.00)	24 (192)	−.18**	[−.30, .05]	22.11 (.00)
PPI-R-VR	20 (–)	.04	[−.05, .14]	36.27 (.45)**	20 (–)	.03	[−.04, .11]	25.48 (.25)	20 (–)	−.03	[−.17, .12]	26.57 (.28)
PICTS-Df	4 (26)	−.38***	[−.54, −.19]	2.04 (.00)	4 (14)	−.23**	[−.39, −.07]	2.84 (.00)	4 (34)	−.48***	[−.59, −.21]	0.72 (.00)
MCSDS	– (–)	—	—	—	2 (–)	−.14	[−.38, .10]	0.54 (.00)	2 (–)	−.32	[−.65, .12]	0.41 (.00)

Note. k = number of effect sizes; k_0 = Orwin's fail-safe N; r_w = weighted mean effect size; CI = confidence interval; Q/I^2 = statistics for test of homogeneity (significant test indicates rejection of hypothesis of homogeneity); PPI/R = Psychopathic Personality Inventory / Psychopathic Personality Inventory – Revised; LSRP = Levenson's Self-Report Psychopathy scales; PAI-PIM = Personality Assessment Inventory – Positive Impression Management; PPI-UV = Psychopathic Personality Inventory – Unlikely Virtues; PPI-R-VR = Psychopathic Personality Inventory-Revised-Virtuous Responding; PICTS-Df = Psychological Inventory of Criminal Thinking Styles – Defensiveness scale; MCSDS = Marlow-Crowne Social Desirability Scale.
* $p < .05$. ** $p < .01$. *** $p < .001$.

gender (Total: $Q_B = .16$, $p = .69$; F1: $Q_B = 1.10$, $p = .30$; F2: $Q_B = 1.28$, $p = .26$), sample type (Total: $Q_B = 3.17$, $p = .08$; F1: $Q_B = .82$, $p = .37$; F2: $Q_B = 1.93$, $p = .16$), or study location (Total: $Q_B = 2.65$, $p = .10$; F1: $Q_B = .10$, $p = .75$; F2: $Q_B = 2.63$, $p = .10$).

It is also important to note that publication bias cannot be ruled out for the significant effect sizes found for some of the subgroups based on the fail-safe *N*'s (k_0 ranges from 14 – 200). For example, the fail-safe *N* is exceptionally small for studies using the PAI-PIM and, thus, reduces confidence that the significant effect found is not spurious. Similar caveats apply for studies using the PICTS-Df, as well as the significant association found between F1 scores and faking good for studies using the LSRP.

Faking Bad

Characteristics for studies included in the meta-analysis for psychopathy measures and indices of faking bad, along with the correlations between the measures of interest, are reported in Appendix 2. Similar to results for faking good, outliers were removed (effect sizes that were considered outliers for faking bad are identified in Appendix 2) and no considerable changes in the I^2 statistic were observed; thus, outliers were retained to maximize the number of effect sizes.

Mean Effect Sizes

Table 3 presents mean effect sizes for the relationship between psychopathy trait measures and faking bad indices for Total, F1, and F2 psychopathy scores based on random effects models. A significant positive association was found between Total and F2 psychopathy and faking bad; F1 was not significantly related to faking bad. Similar to mean effect sizes found for faking good, the fail-safe *N*'s calculated for total and F2 scores ($k_0 = 243$ and 385, respectively) suggest that publication bias is unlikely to be an issue. The Q and I^2 statistics indicated significant heterogeneity among values used to calculate the mean effect sizes for Total, F1, and F2 correlations with faking bad measures; therefore, moderation analyses were performed to examine the sources of this heterogeneity.

Moderation Analyses

Table 4 presents results of moderation analyses (random effects models) for faking bad. Study location moderated the relationship between faking bad and both Total ($Q_b = 6.09$, $p < .05$) and F1 ($Q_b = 4.55$, $p < .05$) scores, but not for F2 ($Q_b = 3.40$, $p = .07$). Specifically, for Total and F1 scores the association with faking bad was significantly higher for U.S. samples than for non-U.S. samples (see Table 4). Tests of homogeneity did not reveal further variability among effect sizes within the remaining Study location × Psychopathy factor cells.

Psychopathy measure was also a significant moderator of the relationship between F1 psychopathy and faking bad ($Q_b = 9.05$, $p < .01$); specifically, the association between F1 and faking bad was significantly stronger for LSRP Primary ($r_w = .22$, $p < .001$) than for PPI-I ($r_w = .00$, $p = .95$). Psychopathy scale did not moderate the relationship between faking bad and F2 ($Q_b = .06$, $p = .81$) or Total ($Q_b = .21$, $p = .65$) psychopathy trait scores. Further, all of the effect size subgroups were homogeneous, with the exception of those for LSRP-Total. Analyses examining the association between LSRP-Total and faking bad revealed that study location moderated this relationship ($Q_b = 8.17$, $p < .01$). Specifically, the mean effect size for U.S. samples ($r_w = .40$, $p < .001$) was much larger than that for non-U.S. samples ($r_w = -.12$, $p = .47$). None of the other potential moderators accounted for heterogeneity among effect sizes between LSRP-Total and faking bad.

Validity scale was a significant moderator of the association between F1 and faking bad, with the mean effect size for the PICTS-Cf ($r_w = .38$, $p < .001$) significantly larger *r* than both the PPI-DR ($r_w = -.02$, $p = .66$) and the PPI-R-DR ($r_w = .07$, $p = .18$)—although analyses for the PICTS-Cf were based on a relatively small sample of effect sizes ($k = 4$). Additionally, the effect sizes among each of the validity scale cells for F1 were homogeneous. Validity scale did not significantly moderate the relationship between faking bad and F2 ($Q_b = 5.30$, $p = .15$) or Total ($Q_b = 6.81$, $p = .08$) psychopathy trait scores. No significant moderation effects were found for gender (Total: $Q_b = 0.15$, $p = .70$; F1: $Q_b = 0.41$, $p = .52$; F2: $Q_b = 0.18$, $p = .67$) or sample type (Total: $Q_b = 0.27$, $p = .60$; F1: $Q_b = 0.46$, $p = .50$; F2: $Q_b = 0.03$, $p = .87$).

Finally, some of the significant effect sizes should be interpreted with caution, as the fail-safe *N*'s (k_0 range from 26 – 595) suggest that publication bias could be contributing to significant effect sizes for some subgroups. Specifically, effect sizes for those studies using non-U.S. samples, the LSRP (primarily those for F1 scores), those using the PAI-NIM, and the PICTS-Cf should be interpreted with caution.

Discussion

Assessing psychopathy with self-report has long been viewed with skepticism (Edens, Hart, et al., 2000; Lilienfeld & Fowler,

Table 3
Mean Effect Sizes for Faking Bad

Psychopathy Scale	k (k_0)	r_u	95% CI_u	r_w	95% CI_w	Q (I^2)
Total Psychopathy	51 (638)	.27*	[.00, .50]	.27***	[.20, .34]	944.02 (.95)***
PPI-I/Primary	50 (–)	.06	[−.21, .32]	.07	[−.00, .14]	749.38 (.93)***
PPI-II/Secondary	50 (750)	.32*	[.05, .54]	.32***	[.23, .40]	1221.87 (.96)***

Note. k = number of effect sizes; k_0 = Orwin's fail-safe *N;* r_u = unweighted mean effect size; CI_u = confidence interval for unweighted mean effect size; r_w = weighted mean effect size; CI_w = confidence interval for weighted mean effect size; Q/I^2 = statistics for test of homogeneity (significant test indicates rejection of hypothesis of homogeneity).
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 4
Moderation Analysis for Faking Bad

	Total				PPI-I/Primary				PPI-II/Secondary			
Moderator	k (k_0)	r_w	95% CI	Q_w (I^2)	k (k_0)	r_w	95% CI	Q (I^2)	k (k_0)	r_w	95% CI	Q (I^2)
Location												
U.S.	35 (525)	.32***	[.25, .40]	27.87 (.00)	34 (170)	.12**	[.04, .20]	40.32 (.18)	34 (595)	.37***	[.27, .45]	29.39 (.00)
Non-U.S.	16 (40)	.14*	[.01, .26]	21.67 (.31)	16 (–)	−.05	[−.17, .08]	9.39 (.00)	16 (64)	.20*	[.05, .35]	19.88 (.25)
Psych Scale												
PPI/R	37 (444)	.26***	[.17, .34]	22.80 (.00)	36 (–)	.00	[−.08, .08]	27.55 (.00)	36 (540)	.32***	[.22, .42]	31.86 (.00)
LSRP	14 (126)	.30***	[.16, .42]	26.12 (.50)*	14 (89)	.22***	[.10, .34]	22.19 (.05)	14 (126)	.30***	[.14, .45]	17.01 (.24)
Validity Scale												
PAI-NIM	3 (–)	.29	[−.00, .57]	1.52 (.00)	3 (8)	.28*	[.03, .49]	2.26 (.12)	3 (–)	.18	[−.17, .49]	2.12 (1.18)
PPI-DR	24 (168)	.24***	[.13, .35]	15.40 (.00)	23 (–)	−.02	[−.12, .07]	24.79 (.11)	23 (208)	.29***	[.17, .40]	17.43 (.00)
PPI-R-DR	20 (147)	.25***	[.14, .37]	31.71 (.40)*	20 (–)	.07	[−.03, .17]	22.21 (.14)	20 (187)	.31***	[.18, .42]	28.75 (.34)
PICTS-Cf	4 (43)	.59***	[.34, .84]	.17 (.00)	4 (26)	.38***	[.18, .55]	.18 (.00)	4 (42)	.57***	[.34, .74]	.53 (.00)

Note. k = number of effect sizes; k_0 = Orwin's fail-safe N; r_w = weighted mean effect size; CI = confidence interval; Q/I^2 = statistics for test of homogeneity (significant test indicates rejection of hypothesis of homogeneity); PPI/R = Psychopathic Personality Inventory/Psychopathic Personality Inventory – Revised; LSRP = Levenson's Self-Report Psychopathy scales; PAI-NIM = Personality Assessment Inventory – Negative Impression Management; PPI-DR = Psychopathic Personality Inventory – Deviant Responding; PPI-R-DR = Psychopathic Personality Inventory-Revised-Deviant Responding; PICTS-Cf = Psychological Inventory of Criminal Thinking Styles – Confusion scale.
* $p < .05$. ** $p < .01$. *** $p < .001$.

2006). At least some of this skepticism may be justified: psychopathic individuals tend to lie frequently and easily (Hare, 1991/2003) and to lack insight into the nature and extent of their psychopathology (Cleckley, 1941/1988; but see Miller et al., 2011, for a different view). Moreover, they may not be highly motivated to cooperate with examiners to provide accurate responses. In this meta-analysis, we investigated one of the principal sources of skepticism regarding the use of self-report measures with psychopaths that derives from their well-known propensity toward prevarication, namely, their presumed tendency to distort their questionnaire responses, especially in a socially desirable or undesirable direction (e.g., Edens, Hart, et al., 2000). To examine this issue, we canvassed the published literature on the relation between the two most widely used self-report indices of psychopathy, namely the PPI (and its derivatives) and the LSRP, and various widely used indicators of social desirability and lying, including positive and negative impression management.

Counter to much conventional wisdom in the psychopathy literature, but consistent with scattered suggestions in previous narrative reviews (e.g., Lilienfeld, 1994), we did not find evidence that scores on psychopathy measures were positively associated with measures of social desirability or faking good in research studies. To the contrary, scores tended to be either negatively or negligibly[3] associated with measures of social desirability or faking good, especially for F2 indices. The findings for F1 indices differed depending on the measure, with the association being negative and statistically significant for the LSRP, but slightly positive but nonsignificant for the PPI/R. Similarly, with the exception of PPI/R F1, for which the association was nonexistent, scores on psychopathy factors were positively correlated with measures of negative impression management.

The lack of a significant association between social undesirability measures and PPI F1 is consistent with findings that this factor, in contrast to those on most other widely used psychopathy measures, reflects a largely adaptive component of psychopathy that is tied to emotional resilience and intact self-esteem (Benning et al., 2003). Hence, high scorers on PPI F1 may be veridically reporting that they do not have many socially undesirable things to say about themselves. The negative correlations between F2 measures and social desirability measures can perhaps best be understood in the context of findings that the latter measures are a complex mix of response styles (stylistic variance) and personality traits (substantive variance; see, e.g., Paulhus, 1991). In particular, meta-analytic research demonstrates that social desirability measures, whether they reflect relatively subtle defensiveness or overt impression management, tend to be negatively correlated with measures of negative emotionality/neuroticism and positively correlated with measures of conscientiousness (Ones, Viswesveran, & Reiss, 1996). Hence, the negative association between psychopathy F2 measures and social desirability measures may reflect the substantial saturation of F2 measures with high negative emotionality and low conscientiousness (Derefinko & Lynam, 2006).

Interestingly, our moderation analyses suggest that psychopathy (mainly Total and F1 scores) exhibits significantly higher associations with faking bad among U.S. samples. This finding is surprising given the consistency found in the construct validity of psychopathy measures across cultures (see, e.g., Lilienfeld, 1998). Assuming that it is replicable, there are several possible explanations for this finding. For example, there may be cultural variation in the interpretation of vocabulary on self-report measures that may affect scores on both psychopathy and validity scales. In addition, items on self-report assessments may be more or less normative in different cultures. For instance, items on the MCSDS assessing impression management ask about socially acceptable yet uncommon behaviors, and items assessing defensive responding ask about socially unacceptable yet common practices. All of these behaviors may be more or less accepted/practiced depending

[3] Interpretation of relative magnitude of correlation coefficients is based on Cohen's (1992) rule of thumb (small = 0.10; medium = 0.30; large = 0.50).

on the cultural context. Nevertheless, these explanations must remain conjectural and require further investigation in cross-cultural research.

Our findings should help to allay concerns (e.g., Hart et al., 1994) that the validity of self-report measures of psychopathy is almost inevitably compromised by the propensity of psychopaths to present themselves as better (e.g., more moral, more virtuous) than they are. Our findings underscore those of Miller et al. (2011), who, using an entirely different approach (i.e., comparing self-reports with informant-reports), came to a similar conclusion–namely, that self-report measures of psychopathy are not necessarily invalid because of response bias. Still, for two reasons, our findings do not imply that response distortion is never a potential concern for the use of self-report measures in assessing psychopathy. First, because there is arguably no "ground truth" for accurate responding on measures of psychopathic traits, we cannot exclude the possibility that psychopathic individuals who obtained low scores on social desirability measures nonetheless underreported their negative attributes. Similarly, the association between psychopathy scores and impression management may be nonlinear; for example, individuals with extremely high levels of psychopathy may be more prone to responding in a socially desirable manner. It is therefore conceivable that scores on self-report psychopathy measures still underestimate the levels of some participants' psychopathic characteristics. Second, our findings were based on individuals who completed the PPI and LRSP in research settings with assurances of confidentiality and often anonymity, with no obvious incentives (e.g., consideration in parole evaluations) for response distortion. Hence, our findings do not rule out the possibility that psychopathic individuals in forensic settings in which incentives for either faking good or faking bad are present are more likely to engage in response distortion than are individuals in the studies examined here (e.g., see Edens & Ruiz, 2006).

Therefore, our findings do not imply that there are no grounds for skepticism regarding the exclusive reliance on self-reports to assess psychopathy. Nor do they imply that self-report measures can safely replace more extensive clinical assessments, such as the Psychopathy Checklist-Revised (Hare, 1991/2003), especially in forensic settings. That said, our findings demonstrate that psychopathic individuals are often willing and able to admit to many socially undesirable traits and behaviors on questionnaires and that such individuals are not necessarily prone to rampant lying and/or positive impression management (see also Lilienfeld, 1994; Miller et al., 2011). This finding, coupled with their ease of administration and lack of reliance on corroborative data (e.g., institutional records, criminal files), suggests that self-report assessments may often be an efficient supplement to clinical assesssments for measuring psychopathy in forensic settings.

Our meta-analysis was marked by several limitations. First, we examined only two self-report measures of psychopathy (three if one counts both the PPI and PPI-R), because the literature on the relation between other self-report psychopathy measures (e.g., the SRP and its derivatives; see Williams, Paulhus, & Hare, 2007) and response distortion was too limited to lend itself to a meta-analytic review. Hence, our conclusions may not apply to all self-report measures of psychopathy. Nevertheless, it is worth noting that the SRP tends to be highly correlated with both measures examined here. For example, Gaughan, Miller, Pryor, and Lynam (2009) reported that SRP-III total scores were correlated $r = .78$ and $r = .71$ with total scores on the PPI and LRSP, respectively. Hence, it is plausible that our conclusions extend to the SRP, although further research should corroborate this possibility. Similarly, we did not include studies using nonself-report (e.g., informant, clinical) assessments of psychopathy. Inclusion of such studies could provide a broader scope for the implications of our findings; moveover, such a multimethod approach could increase the generalizability of our results to nonself-report measures. In addition, this would enable an examination of interaction effects (e.g., examining if the validity of self-report psychopathy scales in predicting scores on clinical assessments of psychopathy is dependent on level of social desirability). Second, as noted earlier, we were unable to examine the moderating effect of incentives versus lack of incentives for response distortion, so our findings may apply only to settings in which individuals complete psychopathy measures without concrete incentives for either positive or negative impression management. Further research examining the association between self-reported psychopathy and social desirability under conditions of explicit incentives is clearly warranted. Third, most of our studies derived from North American samples, with some studies derived from samples in Europe (e.g., Belgium, Netherlands). Nevertheless, we did not locate any studies from nonwestern countries, such as those in Asia or Africa. Hence, the extent to which our findings generalize to nonwestern countries is unknown. A final limitation is that we sought out only published studies. Because of this approach, we cannot rule out the possibility that the significant effect sizes found in the current study are due partly to publication bias; in some cases the fail-safe *N* confirmed this notion. The only way to rule the possibility of publication bias is to thoroughly sample and identify "gray" (buried) literature. Thus, even in those cases where the fail-safe *N* was large, we must interpret the effect sizes with caution (Lipsey & Wilson, 2001). Future research should extend analyses to include associations between self-report indices of psychopathy and socially desirable responding in unpublished studies, such as unpublished doctoral dissertations.

These limitations notwithstanding, our results suggest that previous concerns that the validity of self-report psychopathy measures is inevitably compromised by positive impression management may have been overstated. We found that the two major factors of the two most widely used self-report psychopathy measures were either significantly and negatively associated or (in the case of PPI/R F1) nonsignificantly associated with well-validated indices of social desirability and positive impression management, and significantly and positively associated (again, with the exception of PPI/R F1) with indices of negative impression management. Hence, researchers and clinicians who use questionnaires to detect psychopathic traits can rest assured that their psychopathic participants and clients will often be willing to report traits and behaviors that place them in a decidedly negative light.

References

References marked with asterisk include indicate studies included in the meta-analysis.

*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. doi:10.1037/1040-3590.15.3.340

*Benning, S. D., Patrick, C. J., Salekin, R. T., & Leistico, A. R. (2005). Convergent and discriminant validity of psychopathy factors assessed via self-report: A comparison of three instruments. *Assessment, 12,* 270–289. doi:10.1177/1073191105277110

*Berardino, S. D., Meloy, R., Sherman, M., & Jacobs, D. (2005). Validation of the Psychopathic Personality Inventory on a female inmate sample. *Behavioral Sciences & the Law, 23,* 819–836. doi:10.1002/bsl.666

Book, A. S., Holden, R. R., Starzyk, K. B., Wasylkiw, L., & Edwards, M. J. (2006). Psychopathic traits and experimental induced deception in self-report assessment. *Personality and Individual Differences, 41,* 601–608. doi:10.1016/j.paid.2006.02.011

*Brinkley, C. A., Schmidt, S. A., Smith, S. S., & Newmann, J. P. (2001). Construct validation of a self-report psychopathy scale: Does Levenson's Self-Report Psychopathy scale measure the same constructs as Hare's Psychopathy Checklist Revised? *Personality and Individual Differences, 31,* 1021–1038. doi:10.1016/S0191-8869(00)00178-1

*Campbell, M. A., Doucette, N. L., & French, S. (2009). Validity and stability of the Youth Psychopathic Traits Inventory in a nonforensic sample of young adults. *Journal of Personality Assessment, 91,* 584–592. doi:10.1080/00223890903228679

*Carlson, S. R., Thái, S., & Mclarnon, M. E. (2009). Visual P3 amplitude and self-reported psychopathic personality traits: Frontal reduction is associated with self-centered impulsivity. *Psychophysiology, 46,* 100–113. doi:10.1111/j.1469-8986.2008.00756.x

*Chapman, A. L., Gremore, T. M., & Farmer, R. F. (2003). Psychometric analysis of the Psychopathic Personality Inventory (PPI) with female inmates. *Journal of Personality Assessment, 80,* 164–172. doi:10.1207/S15327752JPA8002_05

*Christopher, K., Lutz-Zois, C. J., & Reinhardt, A. R. (2007). Female sexual-offenders: Personality pathology as a mediator of the relationship between childhood sexual abuse history and sexual abuse perpetration against others. *Child Abuse & Neglect, 31,* 871–883. doi:10.1016/j.chiabu.2007.02.006

Cleckley, H. (1988). *The mask of sanity.* St. Louis, MO: Mosby. (original work published 1941)

*Claes, L., Vertommen, S., Soenens, B., Eyskens, A., Rens, E., & Vertommen, H. (2009). Validation of the Psychopathic Personality Inventory among psychiatric inpatients: Sociodemographic, cognitive, and personality correlates. *Journal of Personality Disorders, 23,* 477–493. doi:10.1521/pedi.2009.23.5.477

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159. doi:10.1037/0033-2909.112.1.155

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment, 13,* 171–188. doi:10.1037/1040-3590.13.2.171

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability in dependent of psychopathology. *Journal of Consulting Psychology, 24,* 349–354. doi:10.1037/h0047358

*Denson, T. F., White, A. J., & Warburton, W. A. (2009). Trait displaced aggression and psychopathy differentially moderate the effects of acute alcohol intoxication and rumination on triggered displaced aggression. *Journal of Research in Personality, 43,* 673–681. doi:10.1016/j.jrp.2009.04.003

*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. doi:10.1521/pedi.2006.20.3.261

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. doi: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. doi:10.1037/1040-3590.12.2.132

*Edens, J. F., Poythress, N. G., Lilienfeld, S. O., Patrick, C. J., & Test, A. (2008). Further evidence of the divergent correlates of the Psychopathic Personality Inventory factors: Prediction of institutional misconduct among male prisoners. *Psychological Assessment, 20,* 86–91. doi:10.1037/1040-3590.20.1.86

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. doi:10.1037/1040-3590.18.2.220

*Egan, V., & Angus, S. (2004). Is social dominance a sex specific strategy for infidelity? *Personality and Individual Differences, 36,* 575–586. doi:10.1016/S0191-8869(03)00116-8

Ekman, P. (1993). Telling lies: Clues to deceit in the marketplace, politics, and marriage. New York: WW. Norton and Company.

*Falkenbach, D., Poythress, N. G., Falki, M., & Manchak, S. (2007). Reliability and validity of two self-report measures of psychopathy. *Assessment, 14,* 341–350. doi:10.1177/1073191107305612

Gacono, G. B., Meloy, J. R., Sheppard, K., Speth, E., & Roske, A. (1995). A clinical investigation of malingering and psychopathy in hospitalized NGRI patients. *Bulletin of the American Academy of Psychiatry and the Law, 23,* 387–397.

*Gaughan, E. T., Miller, J. D., Pryor, L. R., & Lynam, D. R. (2009). Comparing two alternative measures of general personality in the assessment of psychopathy: A test of the NEO PI-R and the MPQ. *Journal of Personality, 77,* 965–996. doi:10.1111/j.1467-6494.2009.00571.x

Hare, R. D. (1990). *The Self-Report Psychopathy Scale-II.* Unpublished test, Department of Psychology, University of British Columbia, Vancouver, Canada.

Hare, R. D. (1991/2003). *The Hare Psychopathy Checklist - Revised.* Toronto, Canada: Multi-Health Systems.

Hart, S. D., Hare, R. D., & Forth, A. E. (1994). Psychopathy as a risk marker for violence: Development and validation of a screening version 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.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21,* 1539–1558. doi:10.1002/sim.1186

Jakobwitz, S., & Egan, V. (2006). The dark triad and normal personality traits. *Personality and Individual Differences, 40,* 331–339. doi:10.1016/j.paid.2005.07.006

*Kruh, I. P., Whittemore, K., Arnaut, G. L. Y., Manley, J., Gage, B., & Gagliardi, G. J. (2005). The concurrent validity of the Psychopathic Personality Inventory and its relative association with past violence in a sample of insanity acquittees. *International Journal of Forensic Mental Health, 4,* 135–145.

Kucharski, L. T., Duncan, S., Egan, S. S., & Falkenbach, D. M. (2006). Psychopathy and malingering of psychiatric disorder in criminal defendants. *Behavioral Sciences & the Law, 24,* 633–644. doi:10.1002/bsl.661

*Lalumiere, M. L., & Quinsey, V. L. (1996). Sexual deviance, antisociality, mating effort, and the use of sexually coercive behaviors. *Personality and Individual Differences, 21,* 33–48. doi:10.1016/0191-8869(96)00059-1

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalised population. *Journal of Personality and Social Psychology, 68,* 151–158. doi:10.1037/0022-3514.68.1.151

Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a serf-report measure of psychopathic personality traits in

noncriminal populations. *Journal of Personality Assessment, 66,* 488–524. doi:10.1207/s15327752jpa6603_3

Lilienfeld, S. O., & Fowler, K. A. (2006). The self-report assessment of psychopathy: Problems, pitfalls, and promises. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 107–132). New York, NY: Guilford Press.

Lilienfeld, S. O., & Widows, M. R. (2005). *Psychological Assessment Inventory-Revised (PPI-R).* Lutz, FL: Psychological Assessment Resources.

Lilienfeld, S. O. (1994). Conceptual problems in the assessment of psychopathy. *Clinical Psychology Review, 14,* 17–38. doi:10.1016/0272-7358(94)90046-9

Lilienfeld, S. O. (1998). Methodological advancements and developments in the assessment of psychopathy. *Behaviour Research and Therapy, 36,* 99–125. doi:10.1016/S0005-7967(97)10021-3

*Lilienfeld, S. O. (1999). The relation of the MMPI-2 Pd Harris-Lingoes subscales to psychopathy, psychopathy facets, and antisocial behavior: Implications for clinical practice. *Journal of Clinical Psychology, 55,* 241–255. doi:10.1002/(SICI)1097-4679(199902)55:2<241::AID-JCLP12>3.0.CO;2-I

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis. Applied social research methods series* (Vol. 49). Thousand Oaks, CA: Sage.

Lynam, D. R., Whiteside, S., & Jones, S. (1999). Self-reported psychopathy: A validation study. *Journal of Personality Assessment, 73,* 110–132. doi:10.1207/S15327752JPA730108

MacNeil, B. M., & Holden, R. R. (2006). Psychopathy and the detection of faking on self-report inventories of personality. *Personality and Individual Differences, 41,* 641–651. doi:10.1016/j.paid.2006.03.004

*Mahaffey, K. J., & Marcus, D. K. (2006). Interpersonal perception of psychopathy: A social relations analysis. *Journal of Social and Clinical Psychology, 25,* 53–74. doi:10.1521/jscp.2006.25.1.53

Malterer, M. B., Lilienfeld, S. O., Neumann, C. S., & Newman, J. P. (2010). Concurrent validity of the Psychopathic Personality Inventory with offender and community samples. *Assessment, 17,* 3–15. doi:10.1177/1073191109349743

*Meier, B. P., Sellbom, M., & Wygant, D. B. (2007). Failing to take the moral high ground: Psychopathy and the vertical representation of morality. *Personality and Individual Differences, 43,* 75–767. doi:10.1016/j.paid.2007.02.001

Miller, J. D., Gaughan, E. T., & Pryor, L. R. (2008). The Levenson Self-Report Psychopathy Scale: An examination of the personality traits and disorders associated with the LSRP factors. *Assessment, 15,* 450–463. doi:10.1177/1073191108316888

Miller, J. D., Jones, S. E., & Lynam, D. R. (2011). Psychopathic traits from the perspective of self- and informant-reports. *Journal of Abnormal Psychology,* 120, 758–764. doi:10.1037/a0022477

*Mokros, A., Menner, B., Eisenbarth, H., Alpers, G. W., Lange, K. W., & Osterheider, M. (2008). Diminished cooperativeness of psychopaths in a prisoner's dilemma game yields higher rewards. *Journal of Abnormal Psychology, 117,* 406–413. doi:10.1037/0021-843X.117.2.406

Morey, L. C. (1991). *The Personality Assessment Inventory: Professional manual.* Odessa, FL: Psychological Assessment Resources.

Neumann, C. S., Malterer, M. B., & Newman, J. P. (2008). Factor structure of the Psychopathic Personality Inventory (PPI): Findings from a large incarcerated sample. *Psychological Assessment, 20,* 169–174. doi:10.1037/1040-3590.20.2.169

Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology, 81,* 660–679. doi:10.1037/0021-9010.81.6.660

Orwin, R. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics, 8,* 157–159. doi:10.2307/1164923

Patrick, C. J., Edens, J. F., Poythress, N. G., Lilienfeld, S. O., & Benning, S. D. (2006). Construct validity of the Psychopathic Personality Inventory two-factor model with offenders. *Psychological Assessment, 18,* 204–208. doi:10.1037/1040-3590.18.2.204

Paulhus, D. L. (1984). Two-component model of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598–609. doi:10.1037/0022-3514.46.3.598

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson & P. R. Shaver (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic Press.

Poythress, N. G., Edens, J. E., & Watkins, M. M. (2001). The relationship between psychopathic personality features and malingering symptoms of major mental illness. *Law and Human Behavior, 25,* 567–582. doi:10.1023/A:1012702223004

*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. doi:10.1177/1073191109351715

*Pryor, L. R., Miller, J. D., & Gaughan, E. T. (2009). Testing two alternative pathological personality measures in the assessment of psychopathy. *Journal of Personality Disorders, 23,* 85–100. doi:10.1521/pedi.2009.23.1.85

*Ray, J. V., & Jones, S. (2011). Self-reported psychopathic traits and their relation to intentions to engage in environmental offending. *International Journal of Offender Therapy and Comparative Criminology, 55,* 370–391. doi:10.1177/0306624X10361582

*Ray, J. V., Poythress, N. G., Weir, J. M., Rickelm, A. (2009). Relationships between psychopathy and impulsivity in the domain of self-reported personality features. *Personality and Individual Differences, 46,* 83–87. doi:10.1016/j.paid.2008.09.005

Ray, J. V., Weir, J. W., Poythress, N. G., & Rickelm, A. (2011). Correspondence between the Psychopathic Personality Inventory and the Psychopathic Personality Inventory-Revised: A look at self-reported personality traits. *Criminal Justice and Behavior, 38,* 375–385. doi:10.1177/0093854811398178

Rogers, R., Bagby, M. R., & Dickens, S. E. (1992). *SIRS, Structured Interview of Reported Symptoms: Professional Manual*: Lutz, FL: Psychological Assessment Resources.

Rogers, R., & Cruise, K. (2000). Malingering and deception among psychopaths. In C. B. Gacono (Ed.), *The clinical and forensic assessment of psychopathy: A practitioner's guide* (pp. 269–284). Mahwah, NJ: Erlbaum.

*Sadeh, N., & Verona, E. (2008). Psychopathic personality traits associated with abnormal selective attention and impaired cognitive control. *Neuropsychology, 22,* 669–680. doi:10.1037/a0012692

*Sandoval, A. R., Hancock, D., Poythress, N. G., Edens, J. F., & Lilienfeld, S. O. (2000). Construct validity of the Psychopathic Personality Inventory in a correctional sample. *Journal of Personality Assessment, 74,* 262–281. doi:10.1207/S15327752JPA7402_7

*Sellbom, M., Ben-Porath, Y. S., Lilienfeld, S. O., Patrick, C. J., & Graham, J. R. (2005). Assessing psychopathic personality traits with the MMPI-2. *Journal of Personality Assessment, 85,* 334–343. doi:10.1207/s15327752jpa8503_10

*Sellbom, M., & Verona, E. (2007). Neuropsychological correlates of psychopathic traits in a non-incarcerated sample. *Journal of Research in Personality, 41,* 276–291. doi:10.1016/j.jrp.2006.04.001

*Stanford, M. S., Houston, R. J., & Baldridge, R. M. (2008). Comparison of impulsive and premeditated perpetrators of intimate partner violence. *Behavioral Sciences and the Law, 26,* 709–722. doi:10.1002/bsl.808

Sullivan, E. A., Kosson, D. S. (2006). Ethnic and cultural variations in psychopathy. In: C. J. Patrick (Ed.), *The Handbook of psychopathy* (pp. 437–458). New York, NY: Guilford Press.

Tellegen, A. (1978/1982). *Brief manual of the Multidimensional Personality Questionnaire.* Unpublished manuscript, University of Minnesota.

*Uzieblo, K., Verschuere, B., & Crombez, G. (2007). The Psychopathic Personality Inventory: Construct validity of the two-factor structure.

Personality and Individual Differences, 43, 657–667. doi:10.1016/j.paid.2007.01.008

*Uzieblo, K. Verschuere, B., Van den Bussche, & Crombez. G. (2010). The validity of the Psychopathic Personality Inventory – Revised in a community sample. *Assessment, 17,* 334–346. doi:10.1177/1073191109356544

*Verschuere, B., Crombez, G., De Clercq, A., & Koster, E. H. W. (2005). Psychopathic traits and autonomic responding to concealed information in a prison sample. *Psychophysiology, 42,* 239–245. doi:10.1111/j.1469-8986.2005.00279.x

*Verschuere, B., Crombez, G., Koster, E. H. W., & De Clercq, A. (2007). Antisociality, underarousal and the validity of the concealed information polygraph test. *Biological Psychology, 74,* 309–318. doi:10.1016/j.biopsycho.2006.08.002

*Walters, G. D., Brinkley, C. A., Magaletta, P. R., & Diamond, P. M. (2008). Taxometric analysis of the Levenson Self-Report Psychopathy scale. *Journal of Personality Assessment, 90,* 491–498. doi:10.1080/00223890802248828

Walters, G. D. (1995). The psychological inventory of criminal thinking styles, part I: Reliability and preliminary validity. *Criminal Justice and Behavior, 22,* 307–332. doi:10.1177/0093854895022003008

*Walters, G. D. Felix, C. M., & Reinoehl, R. (2009). Replicability and cross-gender invariance of a two-dimensional modelof antisociality in male and female college students. *Personality and Individual Differences, 46,* 704–708. doi:10.1016/j.paid.2009.01.028

*Warkentin, J. B., & Gidycz, C. A. (2007). The use and acceptance of sexually aggressive tactics in college men. *The Journal of Interpersonal Violence, 22,* 829–850. doi:10.1177/0886260507301793

Williams, K., Paulhus, D., & Hare, R. (2007). Capturing the four-factor structure of psychopathy in college students via self-report. *Journal of Personality Assessment, 88,* 205–219. doi:10.1080/00223890701268074

Wilson, D. B. (2005). Meta-analysis macros for SAS, SPSS, and Stata. Retrieved from http://mason.gmu.edu/~dwilsonb/ma.html.

*Witt, E. A., Donnellan, B. M., & Blonigen, D. M. (2009). Using existing self-report inventories to measure the psychopathic personality traits of fearless dominance and impulsive antisociality. *Journal of Research in Personality, 43,* 1005–1016. doi:10.1016/j.jrp.2009.06.010

*Zolondek, S., Lilienfeld, S. O., Patrick, C. J., & Fowler, K. A. (2006). The interpersonal measure of psychopathy construct and incremental validity in male prisoners. *Assessment, 13,* 470–482. doi:10.1177/1073191106289861

Appendix A

Effect Sizes for Studies Reporting Correlations Between Self-Report Psychopathy Measures and Validity Scales Assessing Social Desirability/Faking Good

Study	Sample characteristics			Validity measure	Correlations		
PPI/PPI-R	Type	Sex	Location		T (*N*)	F1 (*N*)	F2 (*N*)
Berardino et al. (2005)	Prisoner	Female	United States	PPI-UV	−.08 (102)	−.20 (102)	.09 (102)
Chapman et al. (2003)	Prisoner	Female	United States	PPI-UV	−.17 (153)	.04 (153)	−.19 (153)
Claes et al. (2009)	Clinical	Mixed	Europe	PPI-UV	−.34 (395)	.12 (395)	.54 (395)†
Edens et al. (2008)	Prisoner	Male	United States	PPI-UV	−.17 (50)	.01 (50)	−.32 (50)
Edens et al. (2008)	Prisoner	Male	United States	PPI-UV	−.16 (114)	.28 (114)	−.33 (114)
Kruh et al. (2005)	Forensic	Mixed	United States	PPI-UV	.09 (50)	.21 (50)	.00 (50)
Mahaffey & Marcus (2006)	Forensic	Male	United States	PPI-UV	−.12 (65)	.18 (65)	−.31 (65)
Mokros et al. (2008)	Prisoner	Male	Europe	PPI-R-VR	−.36 (24)	.56 (24)†	−.77 (24)†
Poythress et al. (2010)	Prisoner	Mixed	United States	PPI-UV	−.30 (1593)	.21 (1593)	−.47 (1592)
Ray et al. (2009)	Prisoner	Mixed	United States	PPI-UV	−.14 (89)	.15 (89)	−.29 (88)
Ray et al. (2009)	Prisoner	Mixed	United States	PPI-UV	−.09 (83)	.09 (83)	−.16 (83)
Sandoval et al. (2000)	Prisoner	Mixed	United States	PPI-UV	−.16 (96)	.19 (99)	−.30 (99)
Stanford et al. (2008)	Forensic	Male	United States	PAI-PIM	−.07 (115)	−.45 (115)†	.21 (115)
Uzieblo (2007)	Prisoner	Male	Europe	PPI-R-VR	.23 (165)	−.17 (165)	.37 (165)
Verschuere et al. (2005)	Prisoner	Male	Europe	PPI-UV	−.02 (37)	.42 (37)†	−.32 (42)
Verschuere et al. (2007)	Prisoner	Male	Europe	PPI-UV	−.56 (47)†	−.21 (47)	−.62 (47)
Zolondek et al. (2006)	Prisoner	Male	United States	MPQ-UV	NR	.29 (100)	−.20 (100)
Benning et al. (2003)	Community	Male	United States	PPI-UV	−.12 (353)	−.06 (353)	−.26 (353)
Benning et al. (2005)	Undergrad	Mixed	United States	PPI-UV	−.13 (325)	−.17 (325)	.01 (325)
Campbell et al. (2009)	Undergrad	Mixed	Canada	PPI-R-VR	−.22 (212)	.09 (217)	−.38 (212)
Carlson et al. (2009)	Undergrad	Mixed	Canada	PPI-UV	−.19 (106)	−.05 (106)	−.26 (106)
Denson et al. (2009)	Undergrad	Mixed	Australia	PPI-R-VR	.37 (100)	.28 (100)	.40 (100)
Derefinko & Lynam (2006)	Undergrad	Mixed	United States	PPI-UV	−.09 (336)	.03 (336)	.56 (336)†
Falkenbach et al. (2007)	Undergrad	Male	United States	PPI-UV	−.31 (96)	−.06 (96)	.00 (96)
Gaughan et al. (2009)	Undergrad	Mixed	United States	PPI-R-VR	−.01 (233)	.08 (233)	−.08 (233)
Lilienfeld (1999)	Undergrad	Mixed	United States	PPI-UV	−.03 (113)	NR	NR

(*Appendices continue*)

Appendix A (*continued*)

Study	Sample characteristics			Validity measure	Correlations		
PPI/PPI-R	Type	Sex	Location		T (*N*)	F1 (*N*)	F2 (*N*)
Meier et al. (2007)	Community	Mixed	Europe	PPI-UV	−.38 (58)	−.11 (58)	−.49 (58)
Mokros et al. (2008)	Community	Male	Europe	PPI-R-VR	.29 (24)	.17 (24)	−.01 (24)
Pryor et al. (2009)	Undergrad	Mixed	United States	PPI-R-VR	−.19 (229)	.00 (229)	−.26 (229)
Ray & Jones (2011)	Undergrad	Mixed	United States	PPI-R-VR	−.19 (248)	.05 (253)	−.25 (254)
Sadeh & Verona (2008)	Mixed	Male	United States	PPI-UV	−.25 (150)	.11 (150)	−.42 (150)
Sellbom & Verona (2007)	Undergrad	Mixed	United States	PPI-UV	−.19 (95)	.09 (95)	−.31 (95)
Sellbom et al. (2005)	Undergrad	Mixed	United States	PPI-UV	.07 (281)	.19 (281)	−.07 (281)
Uzieblo et al. (2007)	Undergrad	Male	Europe	PPI-R-VR	.19 (167)	−.18 (167)	.37 (167)
Uzieblo et al. (2007)	Undergrad	Female	Europe	PPI-R-VR	.27 (226)	−.04 (226)	.40 (226)
Uzieblo et al. (2010)	Community	Male	Europe	PPI-R-VR	−.40 (422)	−.05 (422)	−.54 (422)
Uzieblo et al. (2010)	Community	Female	Europe	PPI-R-VR	.03 (253)	−.26 (253)	.28 (253)
Witt et al. (2009)	Undergrad	Mixed	United States	PPI-R-VR	.57 (299)[†]	.17 (299)	.05 (299)
Witt et al. (2009)	Undergrad	Mixed	United States	PPI-R-VR	.18 (304)	.20 (304)	.05 (304)
			LSRP				
Brinkley et al. (2001)	Prisoner	Female	United States	PAI-PIM	−.46 (366)	−.20 (366)	−.59 (366)
Christopher et al. (2007)	Prisoner	Female	United States	BIDR	NR	−.05 (141)	−.07 (142)
Poythress et al. (2010)	Prisoner	Mixed	United States	PPI-UV	−.33 (1584)	−.22 (1585)	−.39 (1584)
Walters (2008)	Prisoner	Male	United States	PICTS-Df	−.27 (521)	−.12 (521)	−.39 (521)
Walters (2008)	Prisoner	Male	United States	PICTS-Df	−.23 (291)	−.08 (291)	−.36 (291)
Walters (2008)	Prisoner	Male	United States	PAI-PIM	−.11 (116)	.04 (116)	−.26 (116)
Campbell et al. (2009)	Undergrad	Mixed	Canada	PPI-R-VR	−.31 (212)	−.24 (217)	−.29 (217)
Egan & Angus (2004)	Community	Mixed	Europe	MCSDS	NR	−.24 (84)	−.45 (84)
Falkenbach et al. (2007)	Undergrad	Male	United States	PPI-UV	−.21 (96)	−.29 (96)	.01 (96)
Gaughan et al. (2009)	Undergrad	Mixed	United States	PPI-R-VR	−.13 (233)	.10 (233)	−.15 (233)
Lalumière & Quinsey (1996)	Mixed	Male	Canada	BIDR	.10 (97)	NR	NR
Pryor et al. (2009)	Undergrad	Mixed	United States	PPI-R-VR	−.35 (229)	−.28 (229)	−.36 (229)
Uzieblo et al. (2010)	Community	Male	Europe	PPI-R-VR	.34 (420)	.23 (420)	.37 (420)
Uzieblo et al. (2010)	Community	Female	Europe	PPI-R-VR	.33 (252)	.20 (252)	.38 (252)
Walters (2009)	Undergrad	Male	United States	PICTS-Df	−.52 (208)	−.39 (208)	−.58 (208)
Walters (2009)	Undergrad	Female	United States	PICTS-Df	−.49 (270)	−.32 (270)	−.58 (270)
Warkentin & Gidycz (2007)	Undergrad	Male	United States	MCSDS	−.14 (297)	−.06 (297)	−.19 (297)
Witt et al. (2009)	Undergrad	Mixed	United States	PPI-R-VR	.05 (304)	.13 (304)	−.11 (304)

Note. Numbers in parentheses are *N* values for each study. PPI = Psychopathic Personality Inventory; PPI-R = Psychopathic Personality Inventory – Revised; T = Total psychopathy trait score; F1 = Factor 1 psychopathy trait score; F2 = Factor 2 psychopathy trait score; PPI-UV = Psychopathic Personality Inventory – Unlikely Virtues; PPI-R-VR = Psychopathic Personality Inventory – Revised – Virtuous Responding; MPQ-UV = Multidimensional Personality Questionnaire – Unlikely Virtues; LSRP = Levenson's Self-report Psychopathy scales; PAI-PIM = Personality Assessment Inventory – Positive Impression Management; BIDR = Balanced Inventory of Desirable Responding; MCSDS = Marlowe-Crowne Social Desirability Scale; PICTS-Df = Psychological Inventory of Criminal Thinking Styles – Defensiveness scale; NR = not reported.

(*Appendices continue*)

Appendix B

Effect Sizes for Studies Reporting Correlations Between Self-Report Psychopathy Measures and Validity Scales Assessing Faking Bad

Study	Sample characteristics			Validity measure	Correlations		
PPI/PPI-R	Type	Sex	Location		T (N)	F1 (N)	F2 (N)
Berardino et al. (2005)	Prisoner	Female	United States	PPI-DR	−.09 (102)	−.22 (102)	.11 (102)
Chapman et al. (2003)	Prisoner	Female	United States	PPI-DR	.43 (153)	.29 (153)	.41 (153)
Edens et al. (2008)	Prisoner	Male	United States	PPI-DR	.05 (50)	−.28 (50)	−.35 (50)†
Edens et al. (2008)	Prisoner	Male	United States	PPI-DR	.34 (114)	−.21 (114)	.51 (114)
Kruh et al. (2005)	Forensic	Mixed	United States	PPI-DR	.48 (50)	−.12 (50)	.58 (50)
Mahaffey & Marcus (2006)	Forensic	Male	United States	PPI-DR	.33 (66)	−.10 (66)	.41 (66)
Mokros et al. (2008)	Prisoner	Male	Europe	PPI-R-DR	.32 (24)	−.09 (24)	.21 (24)
Poythress et al. (2010)	Prisoner	Mixed	United States	PPI-DR	.24 (1593)	−.14 (1593)	.34 (1515)
Ray et al. (2009)	Prisoner	Mixed	United States	PPI-DR	.21 (89)	−.05 (89)	.33 (89)
Ray et al. (2009)	Prisoner	Mixed	United States	PPI-DR	.25 (83)	−.11 (83)	.37 (83)
Sandoval et al. (2000)	Prisoner	Mixed	United States	PPI-DR	.17 (97)	−.17 (99)	.32 (99)
Stanford et al. (2008)	Forensic	Male	United States	PAI-NIM	.28 (115)	.47 (115)	−.09 (115)
Uzieblo et al. (2007)	Prisoner	Male	Europe	PPI-R-DR	.31 (164)	.03 (164)	.37 (164)
Verschuere et al. (2005)	Prisoner	Male	Europe	PPI-DR	.36 (37)	−.01 (37)	.41 (37)
Verschuere et al. (2007)	Prisoner	Male	Europe	PPI-DR	.17 (47)	−.26 (47)	.41 (47)
Benning et al. (2003)	Community	Male	United States	PPI-DR	−.18 (353)	.09 (353)	−.04 (353)
Benning et al. (2005)	Undergrad	Mixed	United States	PPI-DR	.07 (325)	.07 (325)	.08 (325)
Campbell et al. (2009)	Undergrad	Mixed	Canada	PPI-R-DR	.36 (212)	−.04 (217)	.54 (212)
Carlson et al. (2009)	Undergrad	Mixed	Canada	PPI-DR	.38 (106)	.05 (106)	.49 (106)
Denson et al. (2009)	Undergrad	Mixed	Australia	PPI-R-DR	.30 (100)	.23 (100)	.21 (100)
Derefinko & Lynam (2006)	Undergrad	Mixed	United States	PPI-DR	.45 (336)	.09 (336)	−.25 (336)
Falkenbach et al. (2007)	Undergrad	Male	United States	PPI-DR	−.04 (96)	−.01 (96)	.09 (96)
Gaughan et al. (2009)	Undergrad	Mixed	United States	PPI-R-DR	.36 (233)	.01 (233)	.48 (233)
Lilienfeld (1999)	Undergrad	Mixed	United States	PPI-DR	.33 (113)	NR	NR
Meier et al. (2007)	Community	Mixed	Europe	PPI-DR	.16 (57)	−.08 (57)	.29 (57)
Mokros et al. (2008)	Community	Male	Europe	PPI-R-DR	.33 (24)	−.20 (24)	.40 (24)
Pryor et al. (2009)	Undergrad	Mixed	United States	PPI-R-DR	.52 (229)	.05 (229)	.62 (229)
Ray & Jones (2011)	Undergrad	Mixed	United States	PPI-R-DR	.46 (248)	−.00 (251)	.55 (253)
Sadeh & Verona (2008)	Mixed	Male	United States	PPI-DR	.28 (150)	−.18 (150)	.45 (150)
Sellbom & Verona (2007)	Undergrad	Mixed	United States	PPI-DR	.50 (95)	.69 (95)†	.59 (95)
Sellbom et al. (2005)	Undergrad	Mixed	United States	PPI-DR	.45 (281)	−.17 (281)	.58 (281)
Uzieblo et al. (2010)	Community	Male	Europe	PPI-R-DR	.09 (422)	−.15 (422)	.27 (422)
Uzieblo et al. (2010)	Community	Female	Europe	PPI-R-DR	−.42 (253)†	−.10 (253)	−.54 (253)†
Uzieblo et al. (2007)	Undergrad	Male	Europe	PPI-R-DR	.17 (167)	−.06 (167)	.29 (167)
Uzieblo et al. (2007)	Undergrad	Female	Europe	PPI-R-DR	.26 (226)	.06 (226)	.35 (226)
Witt et al. (2009)	Undergrad	Mixed	United States	PPI-R-DR	.17 (299)	.16 (299)	.63 (299)
Witt et al. (2009)	Undergrad	Mixed	United States	PPI-R-DR	.54 (304)	.13 (304)	.60 (304)
			LSRP				
Brinkley et al. (2001)	Prisoner	Female	United States	PAI-NIM	.47 (366)	.31 (366)	.48 (366)
Poythress et al. (2010)	Prisoner	Mixed	United States	PPI-DR	.32 (1586)	.27 (1587)	.28 (1586)
Walters (2008)	Prisoner	Male	United States	PICTS-Cf	.50 (521)	.35 (521)	.53 (521)
Walters (2008)	Prisoner	Male	United States	PICTS-Cf	.49 (291)	.36 (291)	.49 (291)
Walters (2008)	Prisoner	Male	United States	PAI-NIM	.06 (116)	.02 (116)	.09 (116)
Campbell et al. (2009)	Undergrad	Mixed	Canada	PPI-R-DR	.46 (212)	.42 (217)	.38 (217)
Falkenbach et al. (2007)	Undergrad	Male	United States	PPI-DR	−.20 (96)	−.26 (96)	.00 (96)
Gaughan et al. (2009)	Undergrad	Mixed	United States	PPI-R-DR	.44 (233)	.37 (233)	.42 (233)
Pryor et al. (2009)	Undergrad	Mixed	United States	PPI-R-DR	.43 (229)	.39 (229)	.37 (229)
Uzieblo et al. (2010)	Community	Male	Europe	PPI-R-DR	−.43 (420)†	−.34 (420)	−.39 (420)†
Uzieblo et al. (2010)	Community	Female	Europe	PPI-R-DR	−.37 (252)†	−.27 (252)	−.37 (252)†
Walters et al. (2009)	Undergrad	Male	United States	PICTS-Cf	.55 (208)	.45 (208)	.56 (208)
Walters et al. (2009)	Undergrad	Female	United States	PICTS-Cf	.58 (270)	.37 (270)	.68 (270)
Witt et al. (2009)	Undergrad	Mixed	United States	PPI-R-DR	.52 (304)	.49 (304)	.38 (304)

Note. Numbers in parentheses are *N* values for each study. PPI = Psychopathic Personality Inventory; PPI-R = Psychopathic Personality Inventory – Revised; T = Total psychopathy trait score; F1 = Factor 1 psychopathy trait score; F2 = Factor 2 psychopathy trait score; PPI-DR = Psychopathic Personality Inventory – Deviant Responding; PPI-R-DR = Psychopathic Personality Inventory – Revised – Deviant Responding; LSRP = Levenson's Self-report Psychopathy scales; PAI-NIM = Personality Assessment Inventory – Negative Impression Management; PICTS-Cf = Psychological Inventory of Criminal Thinking Styles – Confusion scale; NR = not reported.

† Effect size identified as outlier.