Validity of Rorschach Inkblot Scores for Discriminating Psychopaths From Nonpsychopaths in Forensic Populations: A Meta-Analysis

James M. Wood
University of Texas at El Paso

M. Teresa Nezworski
Lackland Air Force Base

Keli Holloway Allen
University of Texas at El Paso

Scott O. Lilienfeld
Emory University

Howard N. Garb
Lackland Air Force Base

Jessica L. Wildermuth
University of Texas at El Paso

Gacono and Meloy (2009) have concluded that the Rorschach Inkblot Test is a sensitive instrument with which to discriminate psychopaths from nonpsychopaths. We examined the association of psychopathy with 37 Rorschach variables in a meta-analytic review of 173 validity coefficients derived from 22 studies comprising 780 forensic participants. All studies included the Hare Psychopathy Checklist or one of its versions (Hare, 1980, 1991, 2003) and Exner’s (2003) Comprehensive System for the Rorschach. Mean validity coefficients of Rorschach variables in the meta-analysis ranged from −.113 to .239, with a median validity of .070 and a mean validity of .062. Psychopathy displayed a significant and medium-sized association with the number of Aggressive Potential responses (weighted mean validity coefficient = .232) and small but significant associations with the Sum of Texture responses, Cooperative Movement = 0, the number of Personal responses, and the Egocentricity Index (weighted mean validity coefficients = .097 to .159). The remaining 32 Rorschach variables were not significantly related to psychopathy. The present findings contradict the view that the Rorschach is a clinically sensitive instrument for discriminating psychopaths from nonpsychopaths.

Keywords: psychopathy, Rorschach, comprehensive system, forensics, meta-analysis

Supplemental material: http://dx.doi.org/10.1037/a0018998.supp

Psychopaths exhibit a broad and chronic pattern of aberrant personality traits, attitudes, and behaviors, such as superficial charm, grandiosity, lack of guilt, callousness, exploitativeness, irresponsibility, poor impulse control, and often, antisocial behaviors (Cleckley, 1941/1982; Hare, 1993). Extensive research (see summaries by Lilienfeld, 1994 and Widiger, 2007) supported the view that the concepts of psychopath and psychopathy partially overlap but are separable from the diagnostic category of antisocial personality disorder (ASPD) as set forth in the Diagnostic and Statistical Manual of the American Psychiatric Association (4th ed., text rev.; DSM–IV–TR; American Psychiatric Association, 2000). In contrast to ASPD, which focuses on overt and easily measured antisocial behaviors, psychopathy is largely a constellation of personality traits.

Psychopaths are more prone to lying than are nonpsychopaths (Hare, 2003), especially when it is in their interests to do so (Rogers et al., 2002). Indeed, some researchers have referred to “the rampant deception among psychopaths” (Rogers et al., 2002, p. 42; see also Rogers & Cruise, 2000). Nevertheless, research with simulated malingering paradigms calls into question whether psychopaths are better liars (i.e., are less likely to be detected) than are nonpsychopaths (Edens, Buffington, & Tomicic, 2000; see also Poythress, Edens, & Watkins, 2001). Even so, psychopaths’ propensity toward deception, whether successful or not, might sometimes lead to heightened levels of response distortion on self-report measures and interviews (Lilienfeld, 1998; Lilienfeld & Fowler, 2006). Given these limitations, many authors are skeptical of using either self-report measures or interviews based largely on self-report to assess psychopathy. As a consequence, some scholars (Gacono & Meloy, 2009; Hartmann, Nørbech, & Grønnerød, 2006; Loving & Lee, 2006) have proposed that the Rorschach Inkblot Test (Rorschach, 1921) can be especially useful for this purpose because its scores may tap more unconscious or implicit processes relevant to psychopathy (e.g., self-concept, unconscious drives). The idea that psychopathic personality traits can be de-
ected with the aid of the Rorschach is more than half a century old. In the 1940s and 1950s, Robert Lindner (1946, 1950), author of *Rebel Without a Cause: The Hypnoanalysis of a Criminal Psychopath* (Lindner, 1944), proposed that several distinctive Rorschach responses are diagnostic of psychopathy. Specifically, Lindner (1946, 1950; see also Lindner, 1943) claimed that seeing a hammer or tomahawk in Card I of the Rorschach or seeing blood in Card III was often indicative of certain forms of psychopathy. Giel (1945) and Schafer (1948) also argued that the Rorschach was useful for identifying psychopaths and their most salient personality features.

Lindner (1946, 1950), Giel (1945), and Shaffer (1948) based their proposals on clinical observation and impressionistic interpretation of data, without formal verification through statistical tests. An advance toward greater scientific rigor came in the 1960s, when John Exner (1969; see also Raychaudhuri & Mulerji, 1971) published a controlled study in which he reported an unusually high and statistically significant frequency of Rorschach Reflection responses in a sample of psychopaths, a finding he interpreted as indicative of underlying narcissism.

The potential association between psychopathy and Rorschach responses has attracted increasing interest during the past 20 years, due to the work of Carl Gacono and J. Reid Meloy, who published a series of empirical articles in the early 1990s (Gacono, 1990; Gacono & Meloy, 1991, 1992; Gacono, Meloy, & Berg, 1992; Gacono, Meloy, & Heaven, 1990; Meloy & Gacono, 1992), culminating in their influential book *The Rorschach Assessment of Aggressive and Psychopathic Personalities* (Gacono & Meloy, 1994). These six articles and the book drew on psychopathy data from a sample of male inmates, all of whom were diagnosed with ASPD. The inmates were incarcerated in California state prisons, a county jail, and a federal correctional facility. The initial research was based on a sample of 33 inmates (Gacono, 1990). However, the sample size increased over the years, so that the findings reported in Gacono and Meloy’s 1994 book were based on data obtained from 82 inmates (including data from the original 33 participants).

The methodology of Gacono and Meloy’s (e.g., Gacono, 1990; Gacono & Meloy, 1991, 1992) articles was straightforward. Inmates were administered the original version or the revised version of the Hare Psychopathy Checklist (PCL; PCL-R; Hare, 1980, 1991), as well as the Comprehensive System for the Rorschach (CSC; Exner, 1986). Inmates whose PCL or PCL-R scores were 30 or higher were classified as severe psychopaths, whereas inmates whose scores were 29 or lower were classified as moderate psychopaths. The Rorschach scores of the severe psychopathic group were then compared with those of the moderate psychopathic group.

Prior writings on psychopathy and the Rorschach had vague or unspecified criteria to identify psychopaths (e.g., Exner, 1969; Lindner, 1943). In contrast, Gacono and Meloy (1994) introduced an important innovation by using the well established Hare PCL and PCL-R to measure psychopathy. The PCL and the PCL-R are designed for trained raters who evaluate an individual’s personality and behavioral history on the basis of a structured interview and a review of relevant legal and correctional records (Hare, 1980, 1991, 2003). Both the PCL and PCL-R have achieved broad scientific acceptance as reliable and valid indicators of psychopathy in clinical and forensic settings (see Hare, 2003, for a comprehensive review), although use of the PCL-R has superseded use of the PCL since the early 1990s. The PCL-R has demonstrated concurrent validity with other measures of psychopathy and discriminant validity from measures of most Axis I and Axis II disorders and is a valid predictor of violent behavior and criminal recidivism (Hare & Neumann, 2007; Lilienfeld & Fowler, 2006; Porter & Woodworth, 2007; Walters, 2003).

Gacono and Meloy’s research yielded several important findings regarding the concurrent validity of Rorschach scores as indicators of psychopathic traits. For example, Gacono et al. (1990) found that severe psychopaths had significantly higher scores on the Egocentricity Index and were more likely to give Rorschach Reflection and Personal responses, in comparison with moderate psychopaths. These findings were interpreted as consistent with the elevated narcissism and grandiosity characteristic of severe psychopathy. Gacono and Meloy (1991) also reported that severe psychopaths produced significantly fewer Texture (Sum T) and Diffuse Shading responses (Sum Y) on the Rorschach than did moderate psychopaths. These findings were interpreted as consistent with the view that severe psychopathy is characterized by a lack of attachment and low anxiety (Cleckley, 1941/1982; but see Schmitt & Newman, 1999).

Such findings led Gacono and Meloy to conclude that “we have validated the use of the Rorschach as a sensitive instrument to discriminate between psychopathic and nonpsychopathic subjects” (Meloy & Gacono, 2000, p. 236). They contended that just as the concept of psychopathy as measured by the PCL-R allows a more refined understanding of criminal behavior than does the DSM category of ASPD, so does the Rorschach allow a more refined understanding of criminal personality than does the PCL-R:

What begins as a gross categorization of chronic antisocial behavior (*DSM–IV*) moves to a determination of the degree of psychopathic disturbance with the PCL-R. It is further refined through the Rorschach to measure the internal structure and dynamics of the particular patient. (Meloy & Gacono, 2000, p. 238)

The work of Gacono, Meloy, and colleagues has been widely cited. For example, a Google Scholar (http://scholar.google.com) search (November 17, 2009) revealed that Gacono and Meloy’s (1994) book has been cited 168 times and that the articles by Gacono et al. (1990) and Gacono and Meloy (1991) have been cited 39 times and 43 times, respectively.

Following the publication of Gacono and Meloy’s work in the early 1990s, numerous authors (including Gacono and colleagues themselves) attempted to replicate the original findings. For example, Cunliffe and Gacono (2005; Cunliffe, 2002) examined the Rorschach scores of 45 female prison inmates but failed to replicate several important findings. For instance, severe psychopaths did not obtain significantly higher scores on the Egocentricity Index or give significantly more Reflection and Personal responses, in comparison with moderate psychopaths. However, Cunliffe and Gacono (2005, p. 530) proposed that these replication failures were consistent with theoretical expectations and were due to female psychopaths’ distinctive “hysterical character style,” which Cunliffe and Gacono argued is different from the character style of male psychopaths.

In other replication studies, Smith, Gacono, and Kaufmann (1997, 1998) and Smith (1995) examined a sample of incarcerated adolescents diagnosed with conduct disorder, whereas Young, Justice, Erdberg, and Gacono (2000) examined a sample of men-
tally ill prison inmates. These studies also failed to replicate many of the findings reported in Gacono and Meloy’s (1994) earlier work. Several additional studies have also yielded equivocal findings regarding the relationship between psychopathy and Rorschach test responses (e.g., Hartmann et al., 2006; Loving & Russell, 2000; Welsh, 1999).

In an early narrative review of this research literature, Wood, Lilienfeld, Garb, and Nezworski (2000) concluded that no Rorschach variable had exhibited a consistent relationship with psychopathy (see also Wood, Nezworski, Lilienfeld, & Garb, 2003). Some later authors also agreed that although a few Rorschach studies reported positive results, “it has been difficult, however, to replicate most of the findings from these studies” (Hartmann et al., 2006, page 294; see also Loving & Lee, 2006).

However, Gacono, Meloy, and their colleagues (Gacono, Evans, & Viglione, 2008; Gacono, Loving, & Bodholdt, 2001; Meloy, 2005; but see Wood, Lilienfeld, Nezworski, & Garb, 2003) have vigorously responded to critics who question the value of the Rorschach for assessing psychopathy. For example, Gacono and his colleagues (2001) argued that some of the negative findings reviewed by Wood and colleagues (2000) were based on dimensional measures of psychopathy, such as continuous scores on the PCL-R and related measures. Gacono et al. (2001) maintained that positive results for Rorschach indicators should emerge only when categorical operationalizations of psychopathy are used.

More recently, Gacono has argued that critics “do not understand how the Rorschach works” (Gacono, Evans, & Viglione, 2008, p. 5; see also Gacono, Gacono, & Evans, 2008). In a recent chapter in the Oxford Handbook of Personality Assessment, Gacono and Meloy (2009, p. 571) have reasserted their conclusions that the test is “ideally suited” to the assessment of psychopathy and “a nomothetically sensitive instrument in discriminating between psychopathic ASPD and nonpsychopathic ASPD subjects.” To aid clinicians, Gacono and Meloy (2009, p. 572) published a list of “abnormal structural characteristics” that can be expected in the “typical psychopathic Rorschach protocol.”

The present article represents an attempt to address the controversy that has arisen between Gacono and Meloy and their critics. Because survey data suggest that 32% of forensic psychologists routinely use the Rorschach in conducting criminal responsibility assessments (Borum & Grisso, 1995), some of which are almost surely relevant to psychopathy, this controversy is of more than academic importance. We combined the results of Gacono and Meloy’s (1994) original work with the results from all subsequent replications. These pooled results were then analyzed with meta-analysis. The aim was to summarize this body of research and identify the Rorschach variables that can validly discriminate psychopaths from nonpsychopaths in forensic populations.

Method

Literature Search

The PsycInfo database of published and unpublished research in psychology and the ProQuest database of American theses and dissertations were searched on September 1, 2008, for all studies whose titles, abstracts, or descriptors included the term Rorschach co-occurring with psychopath or psychopathy. The 105 articles, chapters, and dissertations thus identified were then examined, and from their citation lists, three additional relevant studies were located. This pool of 108 was then examined to identify all studies in English that met the following criteria: (a) the sample consisted of at least 3 adult or adolescent participants from a forensic or prison population; (b) the participants were administered the CS and some version of the Hare PCL; (c) a subsample of these participants were designated as psychopaths (or severe psychopaths) on the basis of PCL scores; (d) this designation was made with a cutoff within 5 points of the 29/30 cutoff used by Gacono and Meloy (1994); (e) the remaining participants in the sample were sorted into a separate subsample, usually designated as non-psychopaths or moderate psychopaths; and (f) summary statistics (e.g., means, frequencies, or proportions) were provided for the psychopathic and nonpsychopathic subsamples for one or more Rorschach variables from the Exner (2003) CS or from the extended aggression scores proposed by Meloy and Gacono (1992). Also included were studies that met criteria (a) and (b); that did not meet criteria (c), (d), (e), and (f); and that reported correlations between the PCL scores and the Rorschach variables just described.

Applying these criteria to the 108 studies, (a) eight case studies, three book reviews, and three errata were excluded because they did not present data, (b) 25 additional theoretical, historical, or practice-oriented pieces or reviews were excluded because they either did not report data or reported only previously published data; (c) 31 additional research articles were excluded because they did not use the PCL or any version of it; (d) two additional research articles were excluded because they did not report Rorschach findings; (e) one additional research article (Kane, 2004) was excluded because it used a cutoff for the PCL (19/20) that was more than 5 points different from the cutoff of 29/30 used by Gacono and Meloy (1994), (f) six additional research articles were excluded because they supplied insufficient information to calculate validity coefficients for the relationship of the Rorschach to psychopathy (Dansie, 2004; Murphy-Peaslee, 1995; Nunez, 1996; Richards & McCamant, 1995; Weizmann-Henelius, 2006; Young, Justice, & Erdberg, 1999); and (g) three research articles were excluded because the groups they studied differed from each other not only in respect to psychopathy level but also in respect to the type of criminal offense committed, so that psychopathy level and type of offense were confounded (Gacono, Meloy, & Bridges, 2000; Huprich, Gacono, Schneider, & Bridges, 2004; Meloy, Gacono, & Kenney, 1994). In addition to the exclusions just described, (h) four research articles were excluded (Gacono, 1989; Gacono, 1990; Gacono & Meloy, 1992; Siegel, 1999) because they did not report relevant findings for the Exner (2003) CS or for the extended aggression variables proposed by Meloy and Gacono (1992).

Most studies selected for inclusion in the meta-analysis used the PCL-R cutoff score of 29/30 to distinguish psychopaths from nonpsychopaths, just as Gacono and Meloy (1994) did. However, three studies were included whose cut points deviated minimally from the general rule: (a) Smith et al. (1997) used a cutoff of 28/29 on the PCL-R to identify psychopaths in a sample of adolescents; (b) Siemsen (1999, p. 72) used a cutoff of 27/28 on the PCL-R to identify psychopaths in a forensic mentally ill sample, citing a private communication from Robert Hare (August, 1998) to justify this choice; and (c) Hartmann et al. (2006) used a cutoff of 18/19 on the Hare Psychopathy Checklist: Screening Version (PCL:SV),
which follows a different scoring scheme than did the PCL-R and, therefore, has different cut points.

Three dissertations (Darcangelo, 1997; Egozi-Profeta, 1999; Ponder, 1999) reported only correlations between the PCL and the Rorschach scores without the use of cut points. As already noted, Gacono et al. (2001) criticized such studies on the grounds that psychopathy should be measured categorically rather than dimensionally. However, to ensure that the present meta-analysis included all potentially relevant data and because most recent research suggests that psychopathy is dimensional (Edens, Marcus, Lilienfeld, & Pouyhress, 2006; Guay, Ruscio, Knight, & Hare, 2007; Marcus, John, & Edens, 2004; Walters et al., 2007; but see Vasey, Kotov, Frick, & Loney, 2005), we decided to include these three dissertations but to conduct follow-up analyses to determine whether their inclusion had a negative impact on the validity coefficients for the Rorschach variables, as the criticisms of Gacono et al. (2001) imply.

With the above criteria, 22 research reports were identified for inclusion in the meta-analysis, including eight journal articles, one book, one book chapter, four dissertations that were eventually published as articles, and eight dissertations that were never published. Several of these reports included the same or overlapping participants. The number of nonoverlapping samples was 11, and the number of nonoverlapping participants was 780. When two or more research reports reported findings for the same Rorschach variable from the same or overlapping participant samples, we included the findings from the report that had the largest sample size. In the online supplemental material, Appendix 1 lists the 22 research reports, with details regarding the participants, the version of the PCL used, the range of PCL scores used to classify participants, and other study characteristics.

Psychopathy Measures

As may be seen in Appendix 1, four studies used the original PCL to measure psychopathy level, thirteen used the PCL-R, one used the PCL:SV (Hart, Cox, & Hare, 1997), and four used the PCL: Youth Version (PCL:YV; Forth, 1995). Each of these four measures has been used to identify psychopaths in published Rorschach studies. For example, in articles appearing in the Journal of Personality Assessment, Gacono et al. (1990) used the original PCL to identify psychopaths, Gacono et al. (1992) used the PCL-R, Loving & Russell (2000) used the PCL:YV, and Hartmann et al. (2006) used the PCL:SV.

These different versions of the Hare PCL instruments appear to be broadly comparable. For example, the PCL:SV correlates highly (r of approximately .80) with the PCL-R and, in most studies, exhibits a comparable factor structure to that of the PCL-R (Hare & Neumann, 2007; Hart, Hare, & Forth, 1994). In addition, the base rates of categorically operationalized psychopathy for the PCL:SV are similar to, although slightly higher than, those of the PCL-R (Hart et al., 1994).

The PCL:YV similarly appears to have a factor structure broadly comparable with that of the PCL-R (Sulekin, 2007). Although we are unaware of any studies simultaneously examining the PCL-R and PCL:YV (because they are intended for different ages, namely, adults and adolescents, respectively), the base rates of psychopathy yielded by the PCL:YV appear to be similar to, although somewhat lower than, those of the PCL-R (Kosson, Cyterski, Steuerwald, Neumann, & Walker-Matthews, 2002; Murr & Cornell, 2002). Nevertheless, because there is no clear expectation that the adult and adolescent rates of categorically operationalized psychopathy should be comparable, these differences are difficult to interpret.

Rorschach Variables Included in the Meta-Analysis

We included two types of Rorschach variables in the meta-analysis: (a) CS variables that have repeatedly been hypothesized or reported to be related to psychopathy in published studies (e.g., Gacono & Meloy, 1994; Hartmann et al., 2006; Loving & Russell, 2000; Young et al., 2000) and (b) three supplementary aggression scores proposed by Meloy and Gacono (1992). Specifically, the following 16 variables from Exner's (2003) CS were included. They are listed with their abbreviations and their standard interpretive meaning in parentheses: (a) Reflection responses (Reflections; narcissistic personality features); (b) Egocentricity Index (EGOI; excessive self-focus); (c) Personal responses (Personals; defensiveness and grandiosity); (d) Texture responses (Sum T; capacity for interpersonal attachment); (e) Diffuse Shading responses (Sum Y; anxiety or hopelessness); (f) Vista responses (Sum V; painful introspection); (g) Cooperative Movement (COP; openness to collaborative interactions with others); (h) Aggressive Movement (AG; inclination to display assertive behavior); (i) Inaccurate Human Movement (M--; peculiar or disturbed thinking); (j), Pure Human responses (Pure H; ability to perceive others as whole, meaningful objects); (k) White Space responses (Space; oppositionality and anger); (l) Form Dimension responses (FD; capacity for introspection); and (m) the Affective Ratio (AfR; interest in affective stimulation).

Also included were the three components related to the ratio of color and form responses (n) Form–Color responses—(FC), (o) Color–Form responses (CF), and (p) Pure Color responses (Pure C)—whose interpretation concerns modulation of emotional experience. Also included were three extended aggression variables introduced by Meloy and Gacono (1992), (q) Aggressive Content (AgContent), (r) Aggressive Past (AgPast), and (s) Aggressive Potential (AgPotential), which have been hypothesized to reflect propensity for violence.

On the basis of hypotheses or findings in published studies (e.g., Gacono & Meloy, 1994; Hartmann et al., 2006; Loving & Russell, 2000; Young et al., 2000), we expected high psychopathy to be associated with higher mean scores on Reflections, the EGOI, Personals, M–, Space, CF + C > FC + 1, CF, Pure C, AgContent, AgPast, and AgPotential and with lower mean scores on the remaining variables.

Many of the Rorschach variables already listed here have also been analyzed as dichotomous variables in studies on the Rorschach and psychopathy. For example, continuous Reflection responses have frequently been transformed into dichotomous scores (number of Reflections = 0 vs. number of Reflections > 0) and then analyzed in this dichotomized form. Therefore, the following dichotomous Rorschach signs were also included in the meta-analysis: (t) Reflections > 0, (u) EGOI > .44, (v) Personals > 0, (w) Sum T = 0, (x) Sum Y = 0, (y) Sum V > 0, (z) COP = 0, (aa) AG = 0, (bb) M– > 0, (cc) Pure H < 2, (dd) Space > 2, (ee) FD > 0, (ff) AfR < .50, (gg) (CF + C) > FC + 1, (hh) Pure C > 0, (ii) AgContent > 0, (jj) AgPast > 0, and (kk) AgPotential > 0.
Only two of these dichotomous signs were expected to be less common among psychopaths than among nonpsychopaths: Sum \( V > 0 \) and \( FD > 0 \). All the remaining signs were expected to be more common among psychopaths than among nonpsychopaths.

Gacono and Meloy (1994) also hypothesized that some ratios of Rorschach variables (e.g., the ratio of Whole responses to Human Movement responses; \( W:W \)) can distinguish psychopaths from nonpsychopaths. However, an examination of relevant studies revealed that they generally neglected to provide sufficient information to either test these hypotheses statistically or to use them in a meta-analysis. For example, regarding \( W:M \), only one study (Smith, 1995; see also Smith et al., 1997) actually reported the calculated ratio \( W:M \) for each participant, reported this variable’s mean and standard deviation, and reported the statistically tested difference between psychopaths and nonpsychopaths (no significant difference was found). All other studies, including the book by Gacono and Meloy (1994), adopted an alternative approach: Specifically, they calculated the group means of \( W \) and \( M \) and then reported the ratio of these two means as if it were the group mean of \( W:M \). However, this alternative approach is mathematically unsound because the ratio of the means of two variables is not necessarily equal to the mean of the ratio of the two variables. Furthermore, this approach is statistically unsatisfactory because it does not yield a standard deviation for \( W:M \). Thus, the numbers reported in these studies for such ratios as \( W:M \) were insufficient for either statistical hypothesis testing or for use in a meta-analysis and are not included here.

**Calculation of Effect Sizes**

The calculation of validity coefficients followed recommendations by Lipsey and Wilson (2001). For dichotomous Rorschach variables, when participants were divided into two groups (e.g., psychopaths versus nonpsychopaths), \( r \) (phi) was calculated for individual studies, converted to the \( z \)-transform of \( r \), and then averaged across studies with the weights \( N - 3 \) to yield a mean validity coefficient. The mean validity coefficient was then converted back to \( r \) from the \( z \)-transform. The weights of \( N - 3 \) are equal to the inverse of the variance of \( z \)-transformed \( r \) and thus give more weight in the meta-analysis to the most precise estimates of the population effect size.

For nondichotomous Rorschach variables, when participants were divided into two groups, \( d \) was first calculated for each study with means and standard deviations, converted to \( r \) point bserial, converted to the \( z \)-transform of \( r \), and then averaged across studies with the weights \( N - 3 \) to yield a mean validity coefficient. The mean validity coefficient was then converted back to \( r \) from the \( z \)-transform.

Two dissertations (Egozi-Profeta, 1999; Ponder, 1999) reported correlations between Rorschach scores and continuous scores on the PCL-R. These correlations were included as validity coefficients in the meta-analysis. A dissertation by Darcangelo (1997) reported partial correlations between Rorschach scores and continuous scores on the PCL-R, controlling for the number of Rorschach responses and participants’ IQ. These partial correlations were included as validity coefficients in the meta-analysis.

In some research reports, the nonpsychopaths were sometimes subdivided into two separate subgroups with medium versus low scores on the PCL-R. In such cases, for dichotomous Rorschach variables, the results from the medium and low subgroups were combined into a single nonpsychopathic group before calculating \( r \). For nondichotomous Rorschach variables, two initial \( ds \) were calculated by comparing the psychopaths first with the medium subgroup and then with the low subgroup. Then, the weighted average of these two initial \( ds \) was calculated with weights equal to the size of the nonpsychopathic subgroup (medium or low) that had been used to calculate each initial \( d \) respectively.

Only one study (Muntz, 1999) reported findings for the dichotomous variable \( CF + C > FC + 1 \), but one additional study by Welsh (1999) reported findings for the dichotomous variable \( CF + C > FC \). The results for these two variables were combined in the meta-analysis.

As already noted, the book by Gacono and Meloy (1994) did not report separate Rorschach findings for psychopaths and nonpsychopaths. Rather, the book reported numbers for a psychopathic group and for a combined group composed of both the psychopathic group and a nonpsychopathic group. However, although separate Rorschach findings were not reported for the nonpsychopathic group, we used formulas provided by Lipsey and Wilson (2001) to calculate the necessary validity coefficients of the Rorschach variables for discriminating between psychopathic groups and nonpsychopathic groups.

**Results**

**Mean Validity Coefficients of Dichotomous Rorschach Signs and Nondichotomous Rorschach Variables**

In the online supplemental material, Appendixes 2 and 3 list the individual validity coefficients for each Rorschach variable in each study in the meta-analysis. Tables 1 and 2 report the mean validity coefficients (\( r \)) calculated from these individual validity coefficients with a random effects model (Hunter & Schmidt, 2000; Lipsey & Wilson, 2001). As can be seen, the mean validity coefficients of Rorschach variables in the meta-analysis ranged from -.113 to .239, with a median validity of .070 and a mean validity of .062 (\( SD = .079 \)). Psychopathy bore a significant medium-sized association with the number of Aggressive Potential responses (weighted average \( r = .232 \)) and a small but significant association with the Sum of Texture responses (weighted average \( r = .159 \)), Cooperative Movement = 0 (weighted average \( r = .137 \)), the number of Personal responses (weighted average \( r = .115 \)), and the Egocentricity Index (weighted average \( r = .097 \)). Validity coefficients for the remaining 32 Rorschach variables were not significantly different from zero.

Homogeneity tests with the homogeneity statistic \( Q \) were also performed for the mean validity coefficients in Tables 1 and 2. All Rorschach variables found to be significantly nonhomogeneous (\( p < .05 \)) are indicated in the tables.

Tables 3 and 4 compare the frequencies of dichotomous Rorschach signs and the weighted means of nondichotomous Rorschach variables in the meta-analysis, with the Exner (2007) norms (see also Exner & Erdberg, 2005) and the international norms of Meyer, Erdberg, and Shaffer (2007). As can be seen, in many instances, the Exner (2007) norms differ substantially from the international norms, yielding much different interpretations of the frequencies and weighted means from the present meta-analysis. For example, as shown in Table 3, the relative fre-
Table 1

Mean Validity of Nondichotomous Rorschach Variables for Distinguishing Between Psychopathic and Nonpsychopathic Groups

<table>
<thead>
<tr>
<th>Rorschach variable</th>
<th>k</th>
<th>n</th>
<th>r</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>AgPast*</td>
<td>5</td>
<td>310</td>
<td>.239</td>
<td>.147</td>
</tr>
<tr>
<td>AgPotential*</td>
<td>5</td>
<td>310</td>
<td>.232</td>
<td>.050</td>
</tr>
<tr>
<td>M–*</td>
<td>3</td>
<td>227</td>
<td>.176</td>
<td>.252</td>
</tr>
<tr>
<td>Sum T*</td>
<td>9</td>
<td>594</td>
<td>.159</td>
<td>.016</td>
</tr>
<tr>
<td>Personals</td>
<td>9</td>
<td>593</td>
<td>.115</td>
<td>.041</td>
</tr>
<tr>
<td>EGOI</td>
<td>8</td>
<td>570</td>
<td>.097</td>
<td>.023</td>
</tr>
<tr>
<td>AgContent*</td>
<td>5</td>
<td>310</td>
<td>.087</td>
<td>.476</td>
</tr>
<tr>
<td>Sum V</td>
<td>6</td>
<td>356</td>
<td>.081</td>
<td>.138</td>
</tr>
<tr>
<td>FC</td>
<td>3</td>
<td>283</td>
<td>.075</td>
<td>.213</td>
</tr>
<tr>
<td>COP</td>
<td>5</td>
<td>296</td>
<td>.070</td>
<td>.289</td>
</tr>
<tr>
<td>CF</td>
<td>2</td>
<td>130</td>
<td>.070</td>
<td>.600</td>
</tr>
<tr>
<td>Reflectionsa</td>
<td>9</td>
<td>609</td>
<td>.066</td>
<td>.354</td>
</tr>
<tr>
<td>Sum Y</td>
<td>7</td>
<td>478</td>
<td>.064</td>
<td>.190</td>
</tr>
<tr>
<td>Afr</td>
<td>4</td>
<td>328</td>
<td>.041</td>
<td>.467</td>
</tr>
<tr>
<td>Pure H</td>
<td>4</td>
<td>351</td>
<td>.030</td>
<td>.576</td>
</tr>
<tr>
<td>Pure C</td>
<td>2</td>
<td>130</td>
<td>.027</td>
<td>.761</td>
</tr>
<tr>
<td>Space</td>
<td>4</td>
<td>349</td>
<td>.018</td>
<td>.752</td>
</tr>
<tr>
<td>FD</td>
<td>4</td>
<td>349</td>
<td>.042</td>
<td>.446</td>
</tr>
<tr>
<td>AG</td>
<td>8</td>
<td>476</td>
<td>.059</td>
<td>.222</td>
</tr>
</tbody>
</table>

Note. Bold type indicates that mean validity is significantly different from zero, k = number of studies contributing to mean validity coefficient; n = number of participants contributing to mean validity coefficient; AgPast = Aggressive Past; AgPotential = Aggressive Potential; M– = Inaccurate Human Movement; Sum T = Texture responses; Personals = Personal responses; EGOI = Egocentricity Index; AgContent = Aggressive Content; Sum V = Vista responses; FC = Form–Color; COP = Cooperative Movement; CF = Color–Form; Reflections = Reflection responses; Sum Y = Diffuse Shading responses; Afr = Affective ratio; Pure H = Pure Human responses; Pure C = Pure Color; Space = White Space responses; FD = Form Dimension responses; AG = Aggressive Movement.

* Indicates nonhomogeneous effect sizes according to Q statistic, p < .05.

The frequency of Reflections > 0 is 12% in the Exner (2007) norms, 25% in the international norms, 26% among psychopaths, and 18% among nonpsychopaths. Thus, when compared with the Exner (2007) norms, psychopaths and nonpsychopaths appear to be narcissistic. But when compared with the International norms, psychopaths appear to have an average level of narcissism, and nonpsychopaths appear to have a somewhat below-average level of narcissism.

Supplementary Analyses Eliminating Studies That Treated PCL Scores as a Continuous Variable

Gacono and his colleagues (2001) criticized studies on psychopathy and the Rorschach that have treated psychopathy as a dimensional rather than a categorical variable. To determine whether the inclusion of three such studies (Darcangelo, 1997; Egozi-Profeta, 1999; Ponder, 1999) in the present meta-analysis caused the validity coefficients for Rorschach scores to be lower than they would otherwise be, these three studies were temporarily deleted from the data set, and the central analyses were run again with a random effects model. In these reanaly

Table 2

Mean Validity of Dichotomous Rorschach Signs for Distinguishing Between Psychopathic and Nonpsychopathic Groups

<table>
<thead>
<tr>
<th>Rorschach sign</th>
<th>k</th>
<th>n</th>
<th>r</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>COP = 0</td>
<td>3</td>
<td>248</td>
<td>.137</td>
<td>.044</td>
</tr>
<tr>
<td>Sum Y &gt; 0°</td>
<td>6</td>
<td>434</td>
<td>.135</td>
<td>.080</td>
</tr>
<tr>
<td>Afr &lt; .50</td>
<td>2</td>
<td>93</td>
<td>.120</td>
<td>.299</td>
</tr>
<tr>
<td>Sum T &gt; 0°</td>
<td>9</td>
<td>647</td>
<td>.116</td>
<td>.108</td>
</tr>
<tr>
<td>M– &gt; 0</td>
<td>1</td>
<td>82</td>
<td>.115</td>
<td>.303</td>
</tr>
<tr>
<td>AgPast &lt; 0°</td>
<td>2</td>
<td>182</td>
<td>.094</td>
<td>.638</td>
</tr>
<tr>
<td>AgPotential &gt; 0</td>
<td>2</td>
<td>182</td>
<td>.094</td>
<td>.578</td>
</tr>
<tr>
<td>Reflections &gt; 0</td>
<td>8</td>
<td>622</td>
<td>.087</td>
<td>.086</td>
</tr>
<tr>
<td>AG = 0</td>
<td>5</td>
<td>393</td>
<td>.068</td>
<td>.215</td>
</tr>
<tr>
<td>Personals &gt; 0°</td>
<td>5</td>
<td>275</td>
<td>.060</td>
<td>.587</td>
</tr>
<tr>
<td>FD &gt; 0</td>
<td>5</td>
<td>394</td>
<td>.034</td>
<td>.637</td>
</tr>
<tr>
<td>Pure C &gt; 0</td>
<td>4</td>
<td>293</td>
<td>.002</td>
<td>.972</td>
</tr>
<tr>
<td>Sum V &gt; 0°</td>
<td>4</td>
<td>241</td>
<td>-.002</td>
<td>.984</td>
</tr>
<tr>
<td>Space &gt; 2</td>
<td>5</td>
<td>398</td>
<td>-.005</td>
<td>.919</td>
</tr>
<tr>
<td>EGOI &lt; .44</td>
<td>4</td>
<td>379</td>
<td>-.029</td>
<td>.583</td>
</tr>
<tr>
<td>AgContent &gt; 0°</td>
<td>2</td>
<td>182</td>
<td>-.056</td>
<td>.460</td>
</tr>
<tr>
<td>(CF + C) &gt; FC + 1</td>
<td>2</td>
<td>203</td>
<td>-.064</td>
<td>.364</td>
</tr>
<tr>
<td>Pure H &lt; 2°</td>
<td>2</td>
<td>219</td>
<td>-.113</td>
<td>.601</td>
</tr>
</tbody>
</table>

Note. Bold indicates mean validity is significantly different from zero, k = number of studies contributing to mean validity coefficient; n = number of participants contributing to mean validity coefficient; AgPast = Aggressive Past; AgPotential = Aggressive Potential; M– = Inaccurate Human Movement; Sum T = Texture responses; Personals = Personal responses; EGOI = Egocentricity Index; AgContent = Aggressive Content; Sum V = Vista responses; FC = Form–Color; COP = Cooperative Movement; CF = Color–Form; Reflections = Reflection responses; Sum Y = Diffuse Shading responses; Afr = Affective ratio; Pure H = Pure Human responses; Pure C = Pure Color; Space = White Space responses; FD = Form Dimension responses; AG = Aggressive Movement.

* Indicates nonhomogeneous effect sizes according to Q statistic, p < .05.
One purpose of the meta-analysis was to identify potentially promising Rorschach variables for future study. Exploratory analyses were therefore carried out to examine the possible moderating influence on validity coefficients of three study characteristics: (a) gender of participants, (b) age of participants (adults versus adolescents), and (c) publication status (published versus unpublished).

As shown in Table 5, significant moderating effects were found for the following variables: (a) Sum T = 0 was a significant predictor of psychopathy scores among adults (7 studies) and men (6 studies) but not among adolescents (2 studies) or women (2 studies); (b) AgPotential > 0 and AgPast > 0 were significant predictors of psychopathy scores among adolescents (1 study) but not among adults (1 study); (c) AgPotential > 0 and AgPast > 0 were significant predictors of psychopathy scores in an unpublished study (1 study) but were not significant predictors in a published study (1 study); (d) FD > 0 and Pure H < 2 were both significant predictors in the opposite direction from what was hypothesized among adolescents, but were not significant predictors among adults, (e) Pure H < 2 was a significant predictor in the opposite direction from what was hypothesized among published studies (1 study) but not unpublished studies (1 study); (f) the sum of Texture responses (nondichotomous) was a significant predictor of psychopathy among men (7 studies) but not among women (2 studies); and (f) the number of Reflections (nondichotomous) was a significant predictor of Psychopathy scores among adolescents (2 studies) but not among adults (5 studies). In addition, (g) men (3 studies) and women (2 studies) differed significantly from each other on Personals > 0, although this Rorschach variable was not a significant predictor of psychopathy in either gender.

### Methodological Features of Published Versus Unpublished Studies

The moderator analyses described in the previous section and in Table 5 indicated that validity coefficients from published studies and validity coefficients from dissertations did not significantly differ from each other, except in four cases. Several descriptive analyses were also carried out with the study characteristics in the online supplemental material Appendix 1 to determine whether the methodological quality of the published studies differed from that of the dissertations.

These analyses revealed that five (50%) of the 10 published studies and three (25%) of the 12 dissertations neglected to report reliabilities for the Rorschach scores in the meta-analysis. As can be seen, published studies were actually more likely to omit this information than dissertations. When interrater reliabilities of Rorschach scores were reported, they were generally high. Median
reliability was .94 for the published studies and .90 for the dissertations. It should, however, be noted that many studies used percentage agreement as a measure of reliability, probably leading to inflated estimates due to the lack of control for chance (base rate agreement).

It was also found that three (30%) of the published studies and four (33%) of the dissertations did not report interrater reliabilities for PCL-R scores. When interrater reliability was reported for PCL-R scores, it was generally high: Median reliability was .94 for the published studies and .94 for dissertations.

In most studies, the PCL-R was administered either by graduate students or by graduate students and other professionals. Specifically, graduate students administered the PCL-R in nine (90%) of the published studies and seven (58%) of the dissertations. In the remaining studies, the PCL-R was exclusively administered by forensic psychologists. Somewhat surprisingly, graduate student administrators were more common in published studies than in dissertations. Few studies reported the training of the PCL-R administrators. Specifically, only one published article and one dissertation reported that the administrators had been trained in professional workshops.

**Discussion**

Some psychologists have argued that the Rorschach provides valuable information missed by self-report measures or interviews (Gacono & Meloy, 1994, 2009; Loving & Lee, 2006; Loving & Russell, 2000). The present meta-analysis examined 37 Rorschach variables that have been hypothesized by Gacono, Meloy, and other researchers to validly discriminate psychopaths from non-psychopaths in forensic settings. Five of these variables exhibited a modest relationship with psychopathy, but the remaining 32 variables did not. We discuss in detail the five variables with positive findings before considering the broader implications of the meta-analysis.

**Positive Findings**

**AgPotential.** The largest significant validity coefficient was obtained for AgPotential ($r = .232$). This is an unexpected finding because Gacono and Meloy (1994; see also Meloy & Gacono, 1992) reported that AgPotential was not significantly related to psychopathy in their original sample. However, subsequent studies have sometimes found a sizeable relationship (Ballard, 2006; Hartmann et al., 2006).

The relatively good performance of AgPotential is consistent with the hypothesis espoused by Aronow and Reznikoff (1976) that Rorschach content variables generally exhibit higher levels of validity than do Rorschach structural variables. AgPotential is scored when the test taker describes an aggressive act that is about to occur (e.g., “In just a second, one of these bears is going to kill the other”). Meloy and Gacono (1992) conjectured that AgPoten-

---

**Table 4**

Frequencies of Dichotomous Rorschach Signs Among Psychopaths and Nonpsychopaths Contributing to Table 1, Including Comparison With Exner and International Norms

<table>
<thead>
<tr>
<th>Rorschach sign</th>
<th>Psychopaths</th>
<th></th>
<th></th>
<th>Nonpsychopaths</th>
<th></th>
<th></th>
<th></th>
<th>Rel. frequency of sign in normative group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency of sign</td>
<td></td>
<td></td>
<td>Total</td>
<td>Number</td>
<td>%</td>
<td>Total</td>
<td>Number</td>
</tr>
<tr>
<td>COP = 0</td>
<td>3</td>
<td>74</td>
<td>42</td>
<td>57</td>
<td>174</td>
<td>75</td>
<td>43</td>
<td>11</td>
</tr>
<tr>
<td>Sum Y = 0</td>
<td>5</td>
<td>144</td>
<td>66</td>
<td>46</td>
<td>230</td>
<td>91</td>
<td>36</td>
<td>42</td>
</tr>
<tr>
<td>Afr &lt; .50</td>
<td>2</td>
<td>36</td>
<td>24</td>
<td>67</td>
<td>57</td>
<td>37</td>
<td>65</td>
<td>24</td>
</tr>
<tr>
<td>Sum T = 0</td>
<td>7</td>
<td>203</td>
<td>164</td>
<td>81</td>
<td>354</td>
<td>245</td>
<td>69</td>
<td>19</td>
</tr>
<tr>
<td>M− &gt; 0</td>
<td>1</td>
<td>33</td>
<td>18</td>
<td>55</td>
<td>49</td>
<td>21</td>
<td>43</td>
<td>18</td>
</tr>
<tr>
<td>AgPast &gt; 0</td>
<td>2</td>
<td>46</td>
<td>28</td>
<td>61</td>
<td>136</td>
<td>63</td>
<td>46</td>
<td>—</td>
</tr>
<tr>
<td>AgPotential &gt; 0</td>
<td>2</td>
<td>46</td>
<td>14</td>
<td>30</td>
<td>136</td>
<td>27</td>
<td>20</td>
<td>—</td>
</tr>
<tr>
<td>Reflections &gt; 0</td>
<td>7</td>
<td>199</td>
<td>52</td>
<td>26</td>
<td>373</td>
<td>69</td>
<td>18</td>
<td>12</td>
</tr>
<tr>
<td>AG = 0</td>
<td>5</td>
<td>111</td>
<td>67</td>
<td>60</td>
<td>282</td>
<td>196</td>
<td>70</td>
<td>44</td>
</tr>
<tr>
<td>Personals &gt; 0</td>
<td>4</td>
<td>89</td>
<td>57</td>
<td>64</td>
<td>146</td>
<td>86</td>
<td>59</td>
<td>61</td>
</tr>
<tr>
<td>FD &gt; 0</td>
<td>5</td>
<td>147</td>
<td>45</td>
<td>31</td>
<td>247</td>
<td>101</td>
<td>41</td>
<td>80</td>
</tr>
<tr>
<td>Pure C &gt; 0</td>
<td>4</td>
<td>101</td>
<td>40</td>
<td>40</td>
<td>192</td>
<td>69</td>
<td>36</td>
<td>14</td>
</tr>
<tr>
<td>Sum V &gt; 0</td>
<td>4</td>
<td>82</td>
<td>28</td>
<td>34</td>
<td>53</td>
<td>159</td>
<td>33</td>
<td>24</td>
</tr>
<tr>
<td>Space &gt; 2</td>
<td>5</td>
<td>131</td>
<td>33</td>
<td>25</td>
<td>267</td>
<td>71</td>
<td>27</td>
<td>38</td>
</tr>
<tr>
<td>EGOI &gt; .44</td>
<td>4</td>
<td>126</td>
<td>36</td>
<td>29</td>
<td>84</td>
<td>253</td>
<td>33</td>
<td>30</td>
</tr>
<tr>
<td>AgContent &gt; 0</td>
<td>2</td>
<td>46</td>
<td>41</td>
<td>89</td>
<td>136</td>
<td>128</td>
<td>94</td>
<td>—</td>
</tr>
<tr>
<td>(CF + C) &gt; FC + 1</td>
<td>2</td>
<td>74</td>
<td>40</td>
<td>54</td>
<td>129</td>
<td>74</td>
<td>57</td>
<td>26</td>
</tr>
<tr>
<td>Pure H &lt; 2</td>
<td>2</td>
<td>78</td>
<td>47</td>
<td>60</td>
<td>141</td>
<td>83</td>
<td>59</td>
<td>17</td>
</tr>
</tbody>
</table>

**Note.** In the Exner column, adult norms are from Exner and Erdberg (2005) and Exner (2007). In the International column, adult international norms are from Meyer, Erdberg, and Shaffer (2007). The dashes signify that Exner and Erdberg (2005) did not provide means for the three scales AgContent, AgPotential, and AgPast. $k$ = number of studies contributing to mean validity coefficients; AgPast = Aggressive Past; AgPotential = Aggressive Potential; M− = Inaccurate Human Movement; Sum T = Texture responses; Personals = Personal responses; EGOI = Egocentricity Index; AgContent = Aggressive Content; Sum V = Vista responses; FC = Form–Color; COP = Cooperative Movement; CF = Color–Form; Reflections = Reflection responses; Sum Y = Diffuse Shading responses; Afr = Affective ratio; Pure H = Pure Human responses; Pure C = Pure Color; Space = White Space responses; FD = Form Dimension responses; AG = Aggressive Movement; Rel. = relative.

*No international norms available. Proportion was calculated from same normative studies and with the same methods as used by Meyer et al. (2007).*  
*No international norms available. Proportion is from a normative study by Pointkowski (2001), with 122 Rorschach protocols from the Shaffer, Erdberg, and Haronian (1999) nonpatient data set.*
tional responses are related to sadism, but research on this issue has yielded negative results (Darcangelo, 1997). Another possibility is that psychopaths have a tendency to perceive ambiguous situations as having violent consequences (see Vitale, Newman, Serin, & Bolt, 2005), a characteristic that may account for the relationship of AgPotential to psychopathy.

When analyzed as a continuous variable, AgPotential exhibited a small-to-moderate correlation with psychopathy. However, when analyzed as a dichotomous variable, AgPotential was not significantly related to psychopathy (r = .094) and did not attain statistical significance. There appears to be a need to identify a different cut point for AgPotential that yields higher validity.

Sum T. The second largest statistically significant validity coefficient observed in the meta-analysis was for Sum T: Higher psychopathy scores were associated with fewer Texture responses (r = .159). However, when Sum T was dichotomized, the absence of a Texture response (i.e., T = 0) was not significantly related to psychopathy (r = .116). Such findings provide mixed support for the hypothesis that a low number of Texture responses in a protocol is related to “an inability or unwillingness to engage in close, genuine interpersonal exchanges” (Loving & Russell, 2000, p. 138).

COP = 0. COP = 0 was the only dichotomous Rorschach variable in the meta-analysis to exhibit a statistically significant relationship with psychopathy (r = .137). COP = 0 is scored when none of the responses in a Rorschach protocol contains content of individuals acting together cooperatively. Thus, the performance of COP = 0 in the meta-analysis provides additional support for the superior validity of Rorschach content variables (Aronow & Reznikoff, 1976).

COP = 0 is usually interpreted as indicating a lack of interest in collaborative interactions with others. Consistent with this interpretation, COP = 0 was found in 57% of psychopaths’ Rorschach protocols, compared with only 43% for nonpsychopathic forensic participants and 42% of nonpatient adults in the sample of the Rorschach international norms (Meyer et al., 2007).

Personal responses. A Personal response on the Rorschach is scored when respondents refer to their own autobiographical experiences (e.g., “This looks like a dog I used to own”). The present meta-analysis showed a small but significant relationship between the number of such responses and the psychopathy scores (r = .115). The dichotomized form of this variable, Personals > 0, exhibited a smaller and nonsignificant relationship to psychopathy (r = .060).

The meaning of these findings is unclear. Gacono et al. (1990) speculated that Personal responses may occur because the person taking the Rorschach views the examiner as omnipotent and grandiosely identifies with this omnipotence: “Through presentation of

<table>
<thead>
<tr>
<th>Table 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Significant Moderators of Validity Coefficient Size</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Sum T = 0</td>
</tr>
<tr>
<td>Adults</td>
</tr>
<tr>
<td>Male participants</td>
</tr>
<tr>
<td>Adolescents</td>
</tr>
<tr>
<td>Females</td>
</tr>
<tr>
<td>AgPotential &gt; 0</td>
</tr>
<tr>
<td>Adults: Published</td>
</tr>
<tr>
<td>Adolescents: Unpublished</td>
</tr>
<tr>
<td>AgPast &gt; 0</td>
</tr>
<tr>
<td>Adults: Published</td>
</tr>
<tr>
<td>Adolescents: Unpublished</td>
</tr>
<tr>
<td>Personals &gt; 0</td>
</tr>
<tr>
<td>Male participants</td>
</tr>
<tr>
<td>Female participants</td>
</tr>
<tr>
<td>FD &gt; 0</td>
</tr>
<tr>
<td>Adults</td>
</tr>
<tr>
<td>Adolescents: Published</td>
</tr>
<tr>
<td>Texture (nondichotomous)</td>
</tr>
<tr>
<td>Male participants</td>
</tr>
<tr>
<td>Female participants</td>
</tr>
<tr>
<td>Reflections (nondichotomous)</td>
</tr>
<tr>
<td>Adults</td>
</tr>
<tr>
<td>Published</td>
</tr>
</tbody>
</table>

Note. AgPast = Aggressive Past; AgPotential = Aggressive Potential; Sum T = Texture responses; Personals = Personal responses; Reflections = Reflection responses; Pure H = Pure Human responses; FD = Form Dimension responses.
the self-referential, overvalued Personal response (a form of omnipotence), the psychopath bolsters his grandiosity by identifying with the perceived omnipotence of the examiner, thereby preventing any feelings of vulnerability or devaluation” (Gacono et al., 1990, p. 275).

Alternatively, Personal responses may reflect the interpersonal style of psychopaths. In their work on psychopaths’ interpersonal behaviors during PCL-R interviews, Kosson, Steuerwald, Forth, and Kirkhart (1997) found that psychopathic prisoners were more likely than were nonpsychopathic prisoners to make personal comments, ignore personal and professional boundaries, and incorporate interviewers into stories.

Reflection responses and the Egocentricity Index. Of all the Rorschach variables examined in the present study, Reflection responses are of exceptional interest because their connection with psychopathy has been studied since the 1960s (Exner, 1969). These responses, which involve imagery of mirrors or reflected objects (e.g., “A bear reflected in a lake”), are thought to relate to “a narcissistic-like feature of personality” (Exner, 2003, p. 449). Therefore, theory predicts that they should be substantially more frequent among psychopaths than among nonpsychopaths because, as Gacono and Meloy (1994, p. 236) argued, “the criminal psychopath represents a severe, aggressive variant of narcissistic disorder.”

The present findings provide little, if any, support for the validity of Reflection responses as a measure of psychopathic narcissism. The correlation of Reflections with psychopathy scores was $r = .066$ when Reflections were treated as a continuous variable, and $r = .087$ when Reflections were treated as a dichotomous variable (i.e., Reflections $> 0$). In both cases, the correlations were small and nonsignificant (see Tables 1 and 2).

The meta-analytic results are particularly revealing when comparisons are made with the new international norms of the Rorschach (Meyer et al., 2007). As shown in Table 4, 25% of nonpatient adults in the international normative group give at least one Reflection response (Reflections $> 0$). However, in the studies in the present meta-analysis, the percentage of psychopaths who gave at least one Reflection response was 26%, virtually identical to that in the normative group. Thus, a Reflection response seems no more likely to appear in the Rorschach protocol of a psychopath than in the protocol of a typical nonpatient adult. However, exploratory moderator analyses (see Table 5) tentatively suggest that Reflection responses may validly predict psychopathy in adolescents, though not in adults, a possibility that should be interpreted cautiously pending confirmation in future studies.

EGOI scores are partly based on Reflection responses. In the present meta-analysis, continuous EGOI scores exhibited a small but statistically significant relationship with psychopathy ($r = .097$), but the relationship of dichotomous EGOI scores to psychopathy was near zero and nonsignificant ($r = -.029$). As with Reflection responses, when compared with the new international norms for the Rorschach (Meyer et al., 2007), psychopaths’ mean EGOI scores (.39) were virtually identical to the mean scores of nonpatient adults (.38). This finding is provocative, as it seems to indicate either that the level of egocentricity among psychopaths is about the same as in the general population (which seems unlikely given previous findings; see Paulhus & Williams, 2002) or that the EGOI is not measuring egocentricity (Nezworski & Wood, 1995).

Null Findings

No significant relationship emerged between the psychopathy scores and the dichotomous and nondichotomous forms of the following Rorschach variables: Reflections, Sum Y, Sum V, Pure H, FD, M--, Pure C, CF, FC, Afr, Space, AgPast, and AgContent. Although these variables have been hypothesized to bear a valid relationship to psychopathy (e.g., Gacono & Meloy, 1994; Gacono & Meloy, 2009; Loving & Lee, 2006; Loving & Russell, 2000), the results of the present meta-analysis do not support such claims. It is, however, possible that additional positive findings may emerge in the future with different methodologies—for example, if psychopathy were measured as a dimensional variable (cf., Gacono et al., 2001). However, the analyses reported in the results do not clearly support the notion that the validity coefficients were either raised or lowered by measuring psychopathy as either a categorical or dimensional variable.

Possible Moderator Effects

Exploratory moderator analyses reported in Table 5 suggest that three Rorschach scores may have greater validity in some forensic populations than in others. Specifically, it appears that Sum T = 0 and Sum T may bear a statistically significant relationship to psychopathy among men, though not among women. In addition, there is evidence, though somewhat weaker, that the mean number of Reflection responses may bear a moderate relationship to psychopathy among adolescents ($r = .304$), though not adults. However, the mean validity coefficient for adolescents was based on results for only two samples, and the result for only one of these samples was statistically significant (Loving & Russell, 2000). Additional studies are necessary to determine whether this result can be replicated.

Additional analyses suggested other possible moderating effects of participants’ age on Rorschach validity. The validity coefficients for AgPast, AgPotential, and Pure H $< 2$ were all significant in an adolescent sample but not in an adult sample. However, firm conclusions cannot be drawn because only one adult sample and one adolescent sample were involved in each of these moderator analyses. Overall, the findings regarding possible moderator effects might stimulate future research but should be treated as tentative.

Limitations of the Present Findings

Several limitations of the present meta-analysis can be noted. First, the meta-analytic findings for some Rorschach scores are based on a small number of studies or only a few hundred participants. For example, the positive findings for COP $> 0$ are based on only three samples with a combined sample of 248 participants. Similarly, the negative findings for AgPast $> 0$, AgPotential $> 0$, and AgContent $> 0$ are based on only two studies with a combined sample of 182.

Because validity coefficients based on only a few hundred participants are relatively unstable, we recommend additional studies of these variables. In addition, we suggest that results for all the variables included in this meta-analysis, not just a select subset, be included in these studies and all future studies on psychopathy and the Rorschach. Even if a study’s confirmatory hypotheses nar-
rowly focus on only a few Rorschach variables, descriptive findings should also be reported for all variables listed in Tables 1 and 2 of this article. In addition, because the effects of dichotomization are unclear, researchers should report both between-groups differences (psychopaths versus nonpsychopaths) and correlations between continuous Rorschach scores and continuous PCL-R scores. In this way, the research base regarding Rorschach scores and psychopathy is likely to expand much more quickly than it has in the past.

A second potential limitation of the meta-analysis concerns its inclusion of unpublished dissertations. As noted in the Method section, 10 of the studies included in the meta-analysis were publications, four were dissertations that were later published, and eight were dissertations that were never published. The inclusion of unpublished studies is recommended by experts in meta-analysis to detect possible publication bias or file drawer effects (Begg, 1994; Lipsy & Wilson, 2001). However, concerns might arise if the dissertations in the present meta-analysis were of lower methodological quality than the published studies.

In fact, however, the available evidence indicates that the dissertations and published studies in the meta-analysis were similar in regards to methodological quality and results. For instance, as reported in the Results section, approximately the same proportion of dissertations and published studies failed to report interrater reliabilities for Rorschach and PCL-R scores. When reliabilities were reported, they were virtually the same for both kinds of studies. In addition, moderator analyses identified only four significant differences between the validity coefficients of dissertations and published studies. In three of these cases (AgPotential > 0, AgPast > 0, Pure H < 2), a significant validity coefficient was found for dissertations but not for published studies. In the fourth case (Reflections), validity was not significant for either dissertations or published studies. Thus, in three of the four cases, dissertations tended to report higher effect sizes than published studies, contradicting the hypothesis that the inclusion of dissertations in the meta-analysis led to an underestimation of validity.

To avoid future concerns about methodological quality, we recommend that studies on the Rorschach and psychopathy consistently report relevant details concerning their procedures. Specifically, future studies should systematically report (a) interrater reliabilities (i.e., intraclass correlation coefficients) for all Rorschach scores and the PCL-R, (b) the professional qualifications and educational level of all scorers, and (c) the training that scorers and administrators have received in the PCL-R.

**Future Directions**

In future studies, researchers may explore whether some dichotomized Rorschach variables will yield higher validities if new cut points are devised. For example, the cut points of EGOI > .44 and Space > 2 may be too stringent. Researchers may determine whether higher validities can consistently be obtained with EGOI > .50 and Space > 3.

Although most individual Rorschach scores fared poorly in the present study, certain configurations of Rorschach scores may bear a stronger relationship to psychopathy. Thus far, researchers have not identified a replicable relationship between any configuration of Rorschach scores and psychopathy. However, this topic may merit examination in the future.

Finally, it is possible that Rorschach scores with little or no individual validity may nevertheless show incremental validity if they are combined with other data to identify psychopathy (see Meyer et al., 2001). In general, a test with little or no validity when used by itself will have little or no incremental validity when evaluated in the context of other assessment information (Garb, 2003). The idea that low-validity scores can thus be transformed into high-validity ones has been criticized as the “alchemist’s fantasy” (Lilienfeld, Wood, & Garb, 2006), referring to the belief of medieval alchemists that they could transmute lead into gold through elaborate processes. However, future researchers should examine the potential incremental validity of configurations of Rorschach scores above and beyond their individual validities for detecting psychopathy.

**Implications for Forensic Practice**

The present findings are substantially, but not entirely, negative regarding the relationship of Rorschach scores and psychopathy. One score—AgPotential—was found to have a significant mean validity greater than .20 (r = .232). Four additional scores were found to have a significant relationship with psychopathy, but the mean validity coefficients were extremely small in magnitude (r < .160). The large majority of Rorschach variables included in this meta-analysis (32 out of 37) failed to discriminate between psychopaths and nonpsychopaths at above-chance levels.

As a comparison, a meta-analysis by Hiller, Rosenthal, Bornstein, Berry, and Brunell-Neuleib (1999) reported a median validity of .29 for a sample of Rorschach scores. As can be seen, all mean validity coefficients in the present meta-analysis fell short of that figure. Thus, the relationship of Rorschach scores to psychopathy appears to be at best weak in both comparative and absolute terms. Overall, the present findings contradict the view that the Rorschach is a clinically useful instrument for discriminating psychopaths from nonpsychopaths in forensic settings.

Our analyses do not exclude the possibility that some Rorschach variables might be valid for detecting certain features of psychopathy, rather than psychopathy assessed as a global variable. Early research on the PCL-R suggested a two-factor model (Harpur, Hare, & Hakstian, 1989), with the first factor assessing the core interpersonal and affective features of psychopathy and the second factor assessing a chronic impulsive and antisocial lifestyle. More recent work suggests that either a three-factor model (Cooke & Michie, 2001) that cleaves the first factor into two factors or a four-factor model (Hare, 2003) may provide better fit. Future research is necessary to determine whether the Rorschach might be valid for detecting specific factors or facets (e.g., narcissism, lack of guilt, dishonesty, poor impulse control) of the broad psychopathy syndrome. Even if this were the case, such findings would necessitate a substantial revision of Gacono and colleagues’ (e.g., Gacono & Meloy, 2009) assertions, which apply to psychopathy as a whole.

It is also possible that the greatest value of the Rorschach in criminal assessments is not to discriminate psychopaths from nonpsychopaths but to provide a richer picture of personality dynamics (Gacono & Meloy, 2009; Loving & Lee, 2006). However, in evaluating this possibility, it should be borne in mind that recent research on the Rorschach and psychopathy has focused almost exclusively on the type of studies in the present meta-
analysis, in which researchers attempt to use Rorschach scores to discriminate psychopaths from nonpsychopaths. There is little or no research evidence that Rorschach scores can add to the usefulness of forensic assessment of psychopaths in other ways, for instance, by predicting violence, recidivism, or responsiveness to treatment. In the absence of such evidence, its use for these purposes is scientifically unwarranted at present.

References

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


Received July 5, 2009
Revision received November 24, 2009
Accepted November 25, 2009