

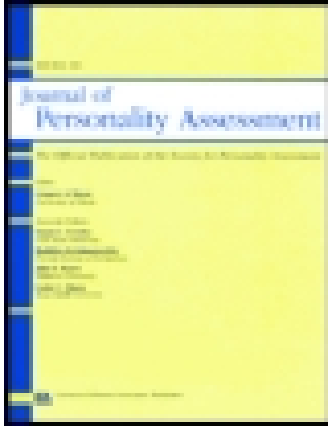
This article was downloaded by: [University of South Florida]

On: 05 December 2014, At: 06:14

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954

Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of Personality Assessment

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/hjpa20>

Coming to Grips With Negative Evidence for the Comprehensive System for the Rorschach: A Comment on Gacono, Loving, and Bodholdt; Ganellen; and Bornstein

James M. Wood , Scott O. Lilienfeld , M. Teresa
Nezworski & Howard N. Garb

Published online: 10 Jun 2010.

To cite this article: James M. Wood , Scott O. Lilienfeld , M. Teresa Nezworski & Howard N. Garb (2001) Coming to Grips With Negative Evidence for the Comprehensive System for the Rorschach: A Comment on Gacono, Loving, and Bodholdt; Ganellen; and Bornstein, Journal of Personality Assessment, 77:1, 48-70, DOI: [10.1207/S15327752JPA7701_04](https://doi.org/10.1207/S15327752JPA7701_04)

To link to this article: http://dx.doi.org/10.1207/S15327752JPA7701_04

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any

losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Coming to Grips With Negative Evidence for the Comprehensive System for the Rorschach: A Comment on Gacono, Loving, and Bodholdt; Ganellen; and Bornstein

James M. Wood

*Department of Psychology
University of Texas at El Paso*

Scott O. Lilienfeld

*Department of Psychology
Emory University*

M. Teresa Nezworski

*School of Human Development
University of Texas at Dallas*

Howard N. Garb

*Pittsburgh V.A. Health Care System
Pittsburgh, Pennsylvania and
Department of Psychiatry,
University of Pittsburgh*

The Comprehensive System (CS; Exner, 1991, 1993) for the Rorschach is currently engulfed in controversy. This comment article responds to 3 articles by Rorschach proponents in this issue of the *Journal of Personality Assessment*. Contrary to the claims of Gacono, Loving, and Bodholdt (this issue), CS scores do not bear a well-demonstrated relationship to psychopathy, antisocial personality disorder, or conduct disorder. Contrary to the claims of Ganellen (this issue), both the original and the revised CS Depression Index (Exner, 1993) bear little or no relationship to depression diagnoses. Furthermore, the scoring reliability of some CS scores is problematic. Although we agree with Bornstein (this issue) that Rorschach scores generally bear little or no relation to psychiatric diagnoses or self-report questionnaires, we believe this

lack of relationship tends to disconfirm hypotheses concerning the validity of the Rorschach. In the spirit of the philosopher Sir Karl Popper, the Rorschach community should not minimize negative evidence or engage in post hoc arguments to immunize the CS against falsification.

The Comprehensive System (CS) for the Rorschach (Exner, 1991, 1993) is currently engulfed in controversy. During the past 2 years, exchanges between critics and proponents of the CS have appeared in three psychological journals (*Assessment* [Archer, 1999], *Journal of Clinical Psychology* [Beutler, 2000], and *Psychological Assessment* [Meyer, 1999]), and additional exchanges are forthcoming in two others (*Clinical Psychology: Science and Practice* [Barlow, 2001], *Journal of Forensic Psychology Practice* [Myers, 2001]). It is fitting, therefore, that disagreements regarding the CS should now receive attention in the journal that began as the *Rorschach Research Exchange* but is now called the *Journal of Personality Assessment*.

At the invitation of this journal's editor, we are writing in response to three articles by Gacono, Loving, and Bodholdt (this issue), Ganellen (this issue), and Bornstein (this issue). Before turning to each article separately, it is necessary to glance back over the past 10 years and consider how perceptions of the Rorschach, and the CS in particular, have changed.

SCORING RELIABILITY OF THE CS

Not long ago, Ritzler (1995) said that "every variable in the Comprehensive System has demonstrated substantial interrater reliability" (p. 230), and that the Rorschach was as reliable as the Wechsler Intelligence scales (Parker, Hanson, & Hunsley, 1988). However, it is becoming increasingly clear that the scoring reliability of many CS scores is considerably lower than was long assumed, and that some CS variables have a level of reliability that is questionable for clinical or forensic work (Acklin, McDowell, Verschell, & Chan, 2000; Gronnerod, 1999; Nakata, 1999/2000; see also Meyer, 1997a, 1997b; Wood & Lilienfeld, 1999; Wood, Nezworski, & Stejskal, 1997). The overall scoring reliability of the CS is markedly lower than that of the Wechsler Adult Intelligence Scale-III (Psychological Corporation, 1997) subscales, which have a median reliability above .95 (using the intraclass correlation coefficient) and a minimum reliability of .90 (Psychological Corporation, 1997). Furthermore, the scoring reliability of the CS when it is used in ordinary clinical practice by ordinary psychologists (i.e., "field reliability") is unknown.

CS NORMS

Not long ago (Weiner, 1998), the CS normative database was described as a "pillar" that supports the scientific status of the Rorschach (p. 55). However, considerable evidence has recently accumulated that the CS norms for many important Rorschach variables are substantially discrepant from the scores of nonpatient children and

adults in the United States and tend to make many normal individuals look emotionally disturbed (Hamel, Shaffer, & Erdberg, 2000; Shaffer, Erdberg, & Haroian, 1999; Wood, Nezworski, Garb, & Lilienfeld, in press). For example, in a recent study of 123 nonpatient adults, Shaffer et al. found that about one sixth of participants scored in the pathological range (≥ 4) on the CS Schizophrenia Index (*SCZI*; Exner, 1993). Over one fourth (29%) gave at least one Reflection response, a supposedly rare Rorschach indicator of narcissism (Exner, 1991, p. 149). Furthermore, the nonpatients often appeared disturbed on Rorschach measures of perceptual accuracy, distorted thinking, and emotional functioning, including Conventional Form ($X + \%$), Distorted Form ($X - \%$), EB style (percentage of Ambitent protocols), *Lambda*, the Affective Ratio, the Form-Color Ratio ($FC: CF + C$), Diffuse Shading Responses (Y), Texture responses (T), the Weighted Sum of Color Responses ($WSumC$), and the Weighted Sum of the 6 Special Scores ($WSum6$).

VALIDITY: RELATIONSHIP OF THE RORSCHACH TO SELF-REPORT MEASURES

Not long ago, researchers (e.g., Archer & Gordon, 1988; Meyer, 1989/1991) predicted that many CS scores should correlate with self-report measures. For example, Meyer (1989/1991, p. 61; but see Meyer, 1992, p. 120) predicted that CS scores should converge with self-report measures of mood and personality. However, because the results of empirical studies have generally failed to support these predictions (e.g. Archer & Krishnamurthy, 1993a, 1993b; Greenwald, 1990, 1991, 1999; Meyer, 1989/1991, 1992, 1993; Nezworski & Wood, 1995), the views of many CS proponents have changed:

Rorschach and MMPI constructs do not converge on a common universe of information in unrestricted heterogeneous samples. This finding is so robust that additional efforts to find cross-method correlates in heterogeneous samples would be redundant. ... Currently, there is enough research to conclude the Rorschach does not consistently or globally measure self-reported characteristics. (Meyer, 1996, p. 572)

VALIDITY: RELATION OF THE RORSCHACH TO DIAGNOSES

Not long ago, it was commonly asserted that CS scores are related to psychiatric diagnoses and could be used for diagnostic purposes (e.g., Weiner, 1997, pp. 10–11). Hundreds of studies examined the hypothesized relationships between Rorschach scores and criterion diagnoses. However, a recent review in the *Journal of Clinical Psychology* by Wood, Lilienfeld, Garb, & Nezworski (2000) concluded

Despite a few positive findings, the Rorschach has demonstrated little validity as a diagnostic tool. Deviant verbalizations and bad form on the Rorschach, and indexes based on these variables, are related to Schizophrenia and perhaps to Bipolar Disorder and Schizotypal Personality Disorder. Patients with Borderline Personality Disorder also seem to give an above-average number of deviant verbalizations. Otherwise the Rorschach has not shown a well-demonstrated relationship to these disorders or to Major Depressive Disorder, Posttraumatic Stress Disorder (PTSD), anxiety disorders other than PTSD, Dissociative Identity Disorder, Dependent, Narcissistic, or Antisocial Personality Disorders [ASPDs], Conduct Disorder [CD], or psychopathy. (p. 395)

Many Rorschach proponents seem to have reached a similar conclusion (Bornstein, this issue). For example, in an apparent change from his earlier opinions (e.g. Weiner, 1997), Weiner (1999) stated

The Rorschach Inkblot Method [RIM] is not a diagnostic test, it was not designed as a diagnostic test, it is not intended to be a diagnostic test, and it does not in fact work very well as a diagnostic test, especially if what is meant by diagnosis is a DSM category. (pp. 336–337)

COMPARABILITY OF THE CS TO OTHER RORSCHACH APPROACHES

Not long ago, it was thought that the psychometric properties of the CS were clearly superior to those of other Rorschach approaches (e.g., Groth-Marnat, 1997; Ritzler, 1995; Viglione, Brager, & Haller, 1991). However, recent studies of the CS norms and CS scoring reliability cast doubt on this view. Furthermore, meta-analyses by Garb, Florio, and Grove (1998) and Hiller, Rosenthal, Bornstein, Berry, and Brunell-Neuleib (1999) found that studies using the CS did not yield significantly higher validity coefficients than other Rorschach approaches. In fact, the estimated validity of the CS in the Hiller et al. meta-analysis (weighted $r = .24$) was almost identical to that of other Rorschach approaches (weighted $r = .25$) when studies were excluded that had inappropriately used comparisons with the problematic CS norms, rather than with genuine comparison groups (Wood et al., in press).¹

As may be seen, assumptions and perceptions regarding the CS have changed in important ways over the past 10 years. In many cases, these shifts have occurred in response to research and reviews by investigators friendly to the CS (e.g., Acklin

¹The corresponding unweighted effect sizes were .27 for Comprehensive System (CS) studies and .22 for non-CS studies. However, virtually all experts on meta-analysis recommend the use of weighted rather than unweighted effect sizes, as discussed by Garb, Wood, Nezworski, Grove, and Stejskal (in press).

et al., 2000; Archer & Krishnamurthy, 1993a, 1993b; Hiller et al., 1999). We have been allocated only limited space for this comment. However, as we respond to the articles by Rorschach proponents, we do our best to identify additional areas in which common assumptions regarding the CS merit reexamination.

RESPONSE TO GACONO ET AL.

The CS and Psychopathy

Although Meloy and Gacono (1995, 2000) claimed that CS scores can discriminate between psychopaths and nonpsychopaths, our (Wood et al., 2000) recent review in the *Journal of Clinical Psychology* arrived at much different conclusions:

Researchers who have attempted to replicate Gacono and Meloy's (1994) Rorschach markers of psychopathy have been almost completely unsuccessful in groups of male, female, and juvenile offenders. ... The scientific evidence does not justify continued use of the Rorschach to identify psychopathy in forensic settings. (p. 414)

In their critique, Gacono et al. (this issue) reject our conclusions and reaffirm the assertions of Meloy and Gacono (2000):

Meloy and Gacono (2000) were correct when they stated ... "We have validated the use of the Rorschach as a sensitive instrument to discriminate between psychopathic and nonpsychopathic (ASPD) subjects" (p. 237). Certain Rorschach variables, in well-designed studies, appear to have some level of discriminative strength in differentiating psychopathic ASPDs from nonpsychopathic ASPDs within the nomothetic paradigm inherent to the necessarily quasi-experimental (i.e., nonrandom assignment to group) designs used in the studies described. (p. 33)

However, Gacono et al. (this issue) do not specify precisely which Rorschach variables possess "discriminative strength," nor do they review empirical literature that consistently supports such optimistic conclusions. In fact, any reasonably impartial review of the research literature leads to directly opposite conclusions from those of Gacono et al.

In the early 1990s, Gacono, Meloy, and their colleagues (e.g., Gacono, 1990; Gacono & Meloy, 1991, 1992; Gacono, Meloy, & Berg, 1992; Gacono, Meloy, & Heaven, 1990) published a series of studies that compared the Rorschach scores of male prisoners who had been classified as "severe psychopaths," "moderate psychopaths," or "non-psychopaths" using the Hare Psychopathy Checklist (PCL; Hare, 1980, 1991). These studies all drew on the same sample of prisoners ($N = 30$ to 43) and therefore did not constitute independent replications. Based on the results from this sample, Gacono and Meloy (1994) concluded that several Rorschach variables

discriminate psychopaths from nonpsychopaths, including Reflection responses, the Egocentricity Index, Texture responses, and Diffuse Shading.

In the middle and late 1990s, 10 published studies and dissertations attempted to replicate the findings of Gacono, Meloy, and their colleagues (e.g., Gacono, 1990; Gacono & Meloy, 1991, 1992; Gacono et al., 1992; Gacono et al., 1990) using various versions of the Hare PCL. These attempted replications included 5 studies of adult male prisoners (Darcangelo, 1996/1997; Egozi-Profeta, 1998/1999; Siemsen, 1999; Welsh, 1999; Young, Justice, Erdberg, & Gacono, 2000), 2 studies of adult female prisoners (Muntz, 1998/1999; Murphy-Peaslee, 1993/1995), 2 studies of male juveniles (Loving & Russell, 2000; Smith, 1994/1995; see also Loving, 1998; Smith, Gacono, & Kaufman, 1997, 1998), and 1 study of mixed gender juvenile offenders (Ponder, 1998/1999).

The results of these replications are overwhelmingly negative. For example, only one study ($N = 66$; Loving & Russell, 2000) found a significant relationship between Reflection responses and psychopathy scores, whereas the nine other studies (total $N = 631$) each found that the relationship was not significant. Two of the studies (total $N = 163$) found mixed support for the relationship of the Egocentricity Index to psychopathy (Smith et al., 1997; Young et al., 2000; but see Smith, 1994/1995), whereas six studies (total $N = 428$) each found that Egocentricity scores were not significantly related to psychopathy (Darcangelo, 1996/1997; Loving & Russell, 2000; Muntz, 1998/1999; Ponder, 1998/1999; Siemsen, 1999; Welsh, 1999). Two studies (total $N = 106$; Darcangelo, 1996/1997; Loving & Russell, 2000) each found that Texture responses were significantly less frequent among psychopaths than nonpsychopaths, but the other eight studies (total $N = 576$) each found that Texture responses were not significantly less frequent among psychopaths. Finally, although six studies (total $N = 395$; Darcangelo, 1996/1997; Egozi-Profeta, 1998/1999; Loving & Russell, 2000; Murphy-Peaslee, 1993/1995; Smith et al., 1997; Welsh, 1999) examined the relationship of Shading responses to psychopathy, none of these attempted replications found a statistically significant relationship between the two variables.

The critique by Gacono et al. (this issue) fails to come to grips with the overwhelmingly negative results that we have just summarized. Instead, the critique suggests that the research results might be different depending on whether the participants were men, women, or juveniles, or whether the PCL was treated as a dichotomous or continuous measure. However, the pattern of findings we have just reported does not support such a contention. The negative results for Reflection responses, the Egocentricity Index, Texture responses, and Shading responses have been found consistently in studies of both men and women, in studies of both adults and juveniles, and in studies using different data analytic approaches. Contrary to the assertions of Gacono et al., the research results are clear: These Rorschach variables bear little or no relationship to psychopathy and are useless for the task of discriminating psychopaths from nonpsychopaths.

The CS, ASPD, and CD

Gacono et al. (this issue) points out that psychopathy is related to but conceptually separable from ASPD. We agree with this long-accepted distinction: In our (Wood et al., 2000, p. 411) recent review of the Rorschach and diagnoses, we discussed the difference between psychopathy and ASPD, and Lilienfeld (one of the authors of this article) has elaborated on the distinction at length (Lilienfeld, 1994, 1998; Lilienfeld & Andrews, 1996). One section of our review was devoted to psychopathy, but a complete separate section was devoted to ASPD and CD.

Just as studies have found no consistent relationship of CS scores with psychopathy, they have similarly found no consistent relationship with ASPD or CD (see review by Wood et al., 2000). A book by Gacono and Meloy (1994) is often cited as evidence that prisoners with ASPD have aberrant scores on the Rorschach. However, the methodology of that book was seriously flawed because scores of prisoners with ASPD (pp. 108–117, 157–169) were compared with the problematic CS norms rather than with true comparison groups (see discussion in Wood et al., in press).

Based on these comparisons with the CS norms, Gacono and Meloy (1994, pp. 108–117, 157–169) concluded that the Rorschachs of prisoners with ASPD showed pathological narcissism (high proportion of Reflection responses), a lack of affectional relatedness (low *Sum T*), impaired interpersonal relationships (low *H*), problems with affect modulation (low ratio of *FC* to *CF + C*), anxiety (high *Sum Y*), and pervasive thought disorder and serious reality-testing problems (low $X + \%$, high $X - \%$, and high *WSum6*).

However, as recent studies (Shaffer et al., 1999; Wood et al., in press) of the CS norms demonstrate, the same pattern of Rorschach scores is shown generally by nonpatient American adults. That is, compared with the CS norms, nonpatient Americans also generally have a high proportion of Reflections, low *Sum T*, low *FC*, high *Sum Y*, low $X + \%$, high $X - \%$, and high *WSum6*. Thus, the most parsimonious interpretation of Gacono and Meloy's (1994) findings seems to be that the problematic CS norms make many adults look pathological, whether those adults are normal Americans or prisoners with ASPD.

Three published studies with comparison groups examined the relationship of CS scores to CD. Weber, Meloy, and Gacono (1992) found that 48 adolescents with a CD diagnosis gave significantly fewer Shading responses, Texture responses, and Pure Human responses than 30 adolescents diagnosed with dysthymia. In contrast, Anderson and Dixon (1993) found that 105 adolescents with CD and 105 adolescents with depressive disorders did not significantly differ in respect to either Shading or Texture responses.² Archer and Krishnamurthy (1997) found that the number

²Although Anderson and Dixon (1993) did not report statistical tests of the between-group differences, we performed *t* tests using the data reported in Table 1 (p. 322) of their article. We found no significant differences for either Shading responses, $t = 0.99$, *ns*; or Texture responses, $t = 1.06$, *ns*.

of Texture responses was not significantly different between 46 adolescents with CD and 106 adolescents with other diagnoses. In fact, Archer and Krishnamurthy (1997) found no significant differences for any of six CS variables with a hypothesized relationship with CD. Furthermore, it is worth noting that two studies of “conduct problems” (rather than formally diagnosed CD) among adolescents found no relationship with CS scores (Karfgin, 1988/1989; Long, 1995).

The study by Weber et al. (1992) used a group of adolescents with a narrowly defined diagnosis of CD, whereas the study by Archer and Krishnamurthy (1997) combined diagnoses of CD, oppositional defiant disorder, and adjustment disorder with disturbance of conduct into one “conduct-disordered” group. Gacono et al. (this issue) criticize our review (Wood et al., 2000) because it compared the results of Weber et al. (1992) with those of Archer and Krishnamurthy (1997):

Although Wood et al. suggested the results of these two studies are contradictory, in fact they may not be because the former found the Rorschach to be useful when CD diagnosis was the inclusion criterion, whereas the latter study found nonsupportive results in a sample defined by more broadly defined criterion. (p. 28)

In response to this criticism, we contacted Robert Archer (of the study by Archer and Krishnamurthy, 1997). He kindly reanalyzed the data from that study by dividing the participants into (a) a group of 37 adolescents with narrowly defined CD (as in the study by Weber et al., 1992), and (b) a group of 115 adolescents with other diagnoses. The results were uniformly negative, just as in their original study. No Rorschach variable showed a significant relationship to CD (R. Archer, personal communication, August 31, 2000). These results provide further support for the conclusions of our review (Wood et al., 2000): No CS variable has shown a consistent, well-replicated relationship to CD.

Continuous Versus Categorical Measurement of Psychopathy

Gacono et al. (this issue) claim that

It is *unrealistic* to expect *the Rorschach or any other instrument* to yield positive findings with respect to assessing the dimensional features of psychopathy unless scores from the PCL-R or PCL:YV (Forth et al., in press) are used as the independent variable to *categorically* distinguish psychopaths from nonpsychopaths or to delineate multiple (i.e., three) levels of severity for comparison [*italics added*]. (p. 29)

In other words, Gacono et al. argue that one should not expect the Rorschach or other measures to correlate with psychopathy if psychopathy is measured as a dimension rather than as a taxon. As a consequence, they maintain that some of the

negative findings concerning the Rorschach and psychopathy that we reviewed can be safely ignored.

This argument is problematic for four reasons. First, although Gacono et al. (this issue) claim that neither the Rorschach nor “any other instrument” should yield positive findings when psychopathy is measured as a dimension, the fact is that numerous studies (e.g., Levenson, Kiehl, & Fitzpatrick, 1995; Lilienfeld & Andrews, 1996; Zagon & Jackson, 1994; for a review, see Lilienfeld, 1998) have reported strong and statistically significant correlations between dimensional scores on psychopathy measures and self-report, peer-report, and interview-based indexes of personality and psychopathology. As these studies show, it is not unrealistic to expect dimensional measures of psychopathy to exhibit substantial correlations with other psychological measures. The negative findings regarding the Rorschach and psychopathy cannot be explained away in this manner.

Second, although Gacono et al. (this issue) assert that psychopathy is taxonic, the research evidence for the taxonicity of psychopathy (e.g., Harris, Rice, & Quinsey, 1994) is questionable on several methodological grounds (see Lilienfeld, 1998). Third, even if psychopathy were taxonic, this would not imply that the cut-off score of 30 on the PCL corresponds to the appropriate hitmax cut—that is, the most likely cutoff score for taxon membership (Meehl & Golden, 1982).

Fourth, despite Gacono et al.’s (this issue) assertion that “the most accurate research findings are likely to result when comparing more extreme scoring participants” (p. 30), it is well known that the use of extreme groups designs tends to result in inflated variances and therefore spuriously large effect sizes (Cronbach, 1960; see also discussion by Wood, Nezworski, Stejskal, Garven, & West, 1999). In other words, Gacono et al. recommend a research approach that is likely to make small relationships appear larger than they really are.

RESPONSE TO GANELLEN

Michael Jordan and the *DEPI*

As we have discussed in prior articles (Garb et al., in press; Wood et al., 2000; Wood, Nezworski, & Stejskal, 1996), most validity studies of the CS Depression Index (*DEPI*) have found that it bears little or no relationship to diagnoses of depression. Ganellen criticizes us because in our discussions we reported studies of the original *DEPI* (Exner, 1986) and studies of the revised *DEPI* (Exner, 1991) together in a single group. He argues that this is a serious mistake, comparable to equating “the current lackluster Chicago Bulls Basketball team (minus Michael Jordan, Scottie Pippen, and Dennis Rodman) with the unbeatable, unstoppable, incomparable Chicago Bulls that won six NBA titles” (Ganellen, this issue, p. 7).

Ganellen’s (this issue) analogy suggests that the original *DEPI* was “lackluster,” whereas the revised *DEPI* is of championship quality, like superathlete Michael Jor-

dan and the old Chicago Bulls. However, the facts do not support this comparison. Eight studies independent of the Rorschach workshops have examined the validity of the original *DEPI*. Seven of these found no significant relationship between the original *DEPI* and psychiatric diagnoses (Archer & Gordon, 1988; Ball, Archer, Gordon, & French, 1991; Carter & Dacey, 1996; Lipovsky, Finch, & Belter, 1989; Sells, 1990/1991; Silberg & Armstrong, 1992; Viglione, Brager, & Haller, 1988), whereas one yielded mixed results (Singer & Brabender, 1993). By comparison, seven studies have examined the revised *DEPI*. Four of these found no significant relationship between the revised *DEPI* and depression diagnoses (Archer & Krishnamurthy, 1997; Ball et al., 1991; Caine, Frueh, & Kinder, 1995; Sells, 1990/1991), two yielded mixed results (Ilonen et al., 1999; Meyer, 1993),³ and only one yielded unmixed positive results (Jansak, 1996/1997; see also the review by Jorgensen, Andersen, & Dam, 2000).

As may be seen, the performance of the original *DEPI* was indeed lackluster. However, the performance of the revised *DEPI* has been at best marginally better and hardly bears comparison with Michael Jordan or other all-star players. Most validity studies of the revised *DEPI* have yielded completely negative results, but a few have not. This pattern of findings suggests that the relation of the revised *DEPI* to depression diagnoses may be non-zero but very low, or that the *DEPI* may be sensitive to some subtypes of depression but not others (Jorgensen et al., 2000). The suggestion that the *DEPI* possesses low validity is supported by an analysis of data from 262 patients reported by Meyer (as cited in Jorgensen et al., 2000, p. 264). Meyer found that the correlation of the revised *DEPI* with depression diagnoses was .111, which was statistically significant using a one-tailed test ($p = .037$). Although non-zero, a correlation of .111 is quite low and much too weak to be useful for distinguishing depressed patients from other diagnostic groups, particularly in settings in which the base rate of depression is considerably lower than 50% (Meehl & Rosen, 1955). Such findings support the circumspect conclusions of Jorgensen et al. (2000): "Until more documentation is available we believe that *DEPI* scores should be interpreted with considerable caution when applied for diagnostic purposes" (p. 278).

In Reference to Archer and Krishnamurthy (1997)

Over 20% of Ganellen's (this issue) text is devoted to criticizing a single study by Archer and Krishnamurthy (1997). In our opinion, the study by Archer and

³Ilonen et al. (1999) found that the *DEPI* distinguished depressed patients from a healthy comparison group, but did not clearly distinguish depressed patients from schizophrenic patients. Meyer's (1993) original article found no significant relationship between *DEPI* scores and depression diagnoses. However, later analyses by Meyer (as reported in a review by Jorgensen, Andersen, & Dam, 2000) with expanded samples indicated a small but significant relationship between *DEPI* scores and depression diagnoses.

TABLE 1
 Relation of MMPI-A scales and the Rorschach *DEPI* to Diagnoses of Depression
 (Corrected Figures for Archer and Krishnamurthy, 1997)

<i>Predictor Variable</i>	<i>Sensitivity</i>	<i>Specificity</i>	<i>Hit Rate</i>
MMPI-A Depression content scale			
$T \geq 60$.57	.75	.68
$T \geq 65$.46	.81	.68
MMPI-A Scale 2			
$T \geq 60$.55	.76	.68
$T \geq 65$.39	.84	.68
Rorschach <i>DEPI</i>			
≥ 4	.66	.46	.53
≥ 5	.36	.71	.58
≥ 6	.07	.89	.59

Note. *DEPI* = Rorschach Depression Index; MMPI-A = Minnesota Multiphasic Personality Inventory-Adolescent.

Krishnamurthy (1997) was well conducted and informative. Every psychologist who uses the Rorschach with adolescents should read it. Its main findings regarding depression were as follows: (a) Scale 2 and the Depression content scale of the MMPI-A (Butcher et al., 1992) were both significantly related to diagnoses of depression in a sample of adolescent inpatients, (b) scores on the revised *DEPI* were not significantly related to diagnoses of depression (although Vista scores were significantly related), and (c) the classification accuracy of the MMPI-A scores was not significantly improved by adding any Rorschach scores.

Ganellen (this issue) advances two main criticisms of the study by Archer and Krishnamurthy (1997). First, Ganellen suggests that in Table 2 of the article by Archer and Krishnamurthy, two sensitivity values for Scale 2 of the MMPI-A have been reversed. Ganellen contends that if the values are rearranged correctly, then Scale 2 is seen to be only slightly more sensitive than the *DEPI*.

In response to this point, we contacted Robert Archer (personal communication, September 28, 2000) of the study by Archer and Krishnamurthy (1997). He informed us that indeed both the sensitivity values and the specificity values had been inadvertently reversed for Scale 2 of the MMPI-A. The corrected information provided by Archer appears in Table 1 of this article. To place the discussion in a broader context, Table 1 includes sensitivity, specificity, and hit rate figures for two MMPI-A scales (the Depression content scale and Scale 2) as well as the *DEPI*. Although Table 1 provides figures for several cutoff scores, we focus our discussion, as Ganellen (this issue) does, on the standard cutoff scores of 65 for the MMPI-A and 5 for the *DEPI*.

As Table 1 shows, Ganellen (this issue) is partially correct: The sensitivity of MMPI-A Scale 2 (.39 at a cutoff of $T \geq 65$) was only slightly higher than the sensi-

tivity of the *DEPI* (.36 at a cutoff score of 5). However, sensitivity provides an incomplete measure of a test's diagnostic performance (Kessel & Zimmerman, 1993) because one can achieve 100% sensitivity simply by classifying all participants as disordered. A fuller picture is provided by also examining the test's specificity and hit rates. As can be seen in Table 1, the specificity of Scale 2 (.84) was higher than that of the *DEPI* (.71). Furthermore, the hit rate for Scale 2 was superior to the hit rate for the *DEPI*. In the study by Archer and Krishnamurthy (1997), a diagnostician could have achieved a hit rate of .50 for depression diagnoses simply by flipping a coin. In contrast, the *DEPI* improved the hit rate to .58 (an increase of .08), and Scale 2 improved it to .68 (an increase of .18). Thus the improvement in the hit rate, above and beyond a coin flip, was more than twice as large with Scale 2 (.18) as with the *DEPI* (.08), although neither measure performed especially well. As Table 1 shows, the performance of the MMPI-A Depression content scale was similar to that of Scale 2.

In summary, although the sensitivity of the three depression measures was similar in the study by Archer and Krishnamurthy (1997), nevertheless the two MMPI-A scales had better specificity and hit rates than the *DEPI*. As a result, the two MMPI-A scales exhibited a significant correlation with depression diagnoses, whereas the *DEPI* did not. Contrary to Ganellen's (this issue) contention, the numbers clearly show that the two MMPI-A scales were related to depression diagnoses, whereas the relation of the *DEPI* to depression diagnoses was weak or negligible.

Ganellen's (this issue) second criticism is that Archer and Krishnamurthy's (1997) sample of adolescent patients may have been "anomalous." To support this assertion, he claims that the mean *Lambda* score of these patients (1.84) was "quite unusual" compared with the scores reported by Exner and Weiner (1995) for adolescents with depression (.74) and CD (.91).⁴

In response, we point out that the *Lambda* scores reported by Archer and Krishnamurthy (1997) were not in fact anomalous. For example, Hamel et al. (2000) reported that in a sample of 100 normal children, ages 6 to 12, the mean *Lambda* was 1.91. This figure is very close to the 1.84 figure reported for the adolescent sample of Archer and Krishnamurthy (1997), although quite discrepant from figures reported by Exner and Weiner (1995) for normal and disturbed children. Similarly, Krishnamurthy and Archer (1997) identified five other studies that, like theirs, had examined children or adolescents with psychiatric disturbances or CD (Abraham, Lepisto, Lewis, Schultz, & Finkelberg, 1994; Anderson & Dixon, 1993; Brinkman, Overholser, & Klier, 1994; Gacono & Meloy, 1994;

⁴Ganellen (this issue) argues that the findings of Archer and Krishnamurthy (1997) are anomalous for two variables, *Lambda* and the number of Pure Form responses. However, these two variables are very closely related. Specifically, if R equals the total number of Rorschach responses, and if F equals the number of Pure Form responses, then $Lambda = F / (R - F)$. Thus, our comments regarding *Lambda* almost certainly apply to the number of Pure Form responses as well.

Lipovsky et al., 1989; see also Cotugno, 1995). In these five studies, with sample sizes ranging from 50 to 210, the lowest mean *Lambda* reported for a psychiatric group or group with CD was 1.33, the highest was 2.50, and the median was 1.60. Again, these findings are very similar to those of Archer and Krishnamurthy (1997).

It is the figures reported by Exner and Weiner (1995), not the figures reported by Archer and Krishnamurthy (1997), that appear anomalous when compared with those of other researchers. As we have already noted, there are serious questions regarding the representativeness of the CS norms for children and adolescents. So too, there may be questions regarding the figures that Exner and Weiner reported for *Lambda* among adolescents with psychiatric disorders or CD.

Scoring Reliability of the CS

In previous articles (Wood & Lilienfeld, 1999; Wood et al., 1996), we have questioned whether the scoring reliability of all CS scores is as uniformly excellent as has often been claimed (Exner, 1993, p. 23; Groth-Marnat, 1997, p. 397). Ganellen (this issue) suggests that our concerns about CS scoring reliability were effectively addressed in a meta-analysis by Meyer (1997a, 1997b).

In response, we wish to point out that Meyer's (1997a) meta-analysis was of 10 Rorschach "segments." In the segment approach, individual Rorschach scores are combined into a single segment and then their interrater reliability is assessed simultaneously in a single omnibus statistical test. In a critique of Meyer's (1997a) meta-analysis (Wood et al., 1997, p. 493; but see Meyer, 1997b), we argued that the segment approach to reliability can "mask inaccuracy." For example, when the 23 separate Determinant scores, such as Vista or Shading, are combined into one "Determinant segment," then the omnibus segment reliability may appear quite good, even though the reliability of some individual determinants may be poor.

A recent study of CS interrater reliability by Acklin et al. (2000) provides support for our critique of segment reliability. For example, Meyer's (1997a) meta-analysis reported that the segment reliability of Determinants was .85 as measured by Cohen's kappa. However, Acklin et al. found that individual Determinants often exhibited substantially lower reliability. For example, the reliability of Vista responses as measured by kappa was .553 in one sample of Acklin et al. The reliability of Shading responses was .698 in one sample and .589 in another. As these numbers show, the reliability of individual Rorschach determinants can be problematic, even though the corresponding segment reliability appears quite good.

Although Ganellen (this issue) cites Meyer's (1997a) questionable meta-analysis of Rorschach segments, he neglects to cite or discuss three recent empirical studies that focused on the interrater reliability of individual Rorschach scores (Acklin et al., 2000; Gronnerod, 1999; Nakata, 1999/2000). Of these three, the study by Acklin et al. (2000) had the strongest design because it used the most pro-

ficient scorers, examined a very broad range of CS variables (nearly 100) instead of just a few, and used the most appropriate reliability statistics (kappa and the intraclass correlation coefficient).

Despite their differences in methodological quality, all three studies (Acklin et al., 2000; Gronnerod, 1999; Nakata, 1999/2000) yielded very similar results: Many CS scores showed excellent scoring reliability, but a substantial proportion did not. Contrary to past claims, about half of CS scores appear to have reliability below .85, and a substantial proportion (9% to 25% in the study by Acklin et al., 2000) apparently fail to meet even the lax standard of reliability greater than .60.

Ganellen (this issue) takes issue with our recommendation that the field reliability of the CS be examined. For example, we suggested (Wood et al., 1997) that the everyday, on-the-job scoring reliability of custody evaluators or school psychologists be studied. Ganellen (this issue) argues that our recommendation is unfair because “this standard has never been applied to any other commonly used psychological assessment instrument . . . such as responses to the Wechsler Adult Intelligence Scale–III or the Wechsler Memory Scale–III” (p. 2).

However, Ganellen’s (this issue) claim is in error. As Hunsley and Bailey (in press) pointed out, “actually, there is a sizeable literature on the accuracy with which the various Wechsler intelligence tests are scored in clinical settings.” Hunsley and Bailey cited relevant studies of the Wechsler tests by Slate, Jones, Murray, and Coulter (1993); Slate, Jones, Coulter, and Covert (1992); Whitten, Slate, Jones, and Shine (1994); and Ryan, Prifitera, and Powers (1983). Contrary to Ganellen’s assertion, the field reliability of the Wechsler tests has been examined repeatedly by researchers. It is time that the field reliability of the CS be examined in a similar way.

RESPONSE TO BORNSTEIN

The (Lack of) Relationship Between the Rorschach and Diagnoses

Bornstein (this issue) argues that “the RIM is not a diagnostic tool. RIM scores should not be expected to correlate highly with *DSM–IV* diagnoses” (p. 45). We are in complete agreement with Bornstein on this point. In our recent review of research on the Rorschach and psychiatric diagnoses (Wood et al., 2000), we arrived at virtually the same conclusions: With a few exceptions, Rorschach scores bear little or no relation to diagnoses. Psychologists who use the Rorschach should neither expect nor claim that the scores have a well-demonstrated relation to psychopathology as defined in the *DSM–IV* (American Psychiatric Association, 1994) or similar manuals.

Although we agree with Bornstein (this issue) that the Rorschach is unrelated to most psychiatric diagnoses, we interpret this fact much differently than he does.

Specifically, whereas he adopts the position that the overwhelmingly negative findings do not raise doubts concerning the validity of the Rorschach, we believe that they do. To understand the difficulty with Bornstein's position, a historical perspective is necessary.

For many years, virtually all experts on the Rorschach claimed that it is a valid measure of psychiatric diagnoses. For example, Exner (1991, 1993) gave names to several CS indexes that implied diagnostic utility (for example, the *SCZI*, *DEPI*, and Obsessive Style Index). Furthermore, he made related assertions regarding validity, including the claim that an elevated score on the *DEPI* "correlates very highly with a diagnosis that emphasizes serious affective problems" (Exner, 1991, p. 146).

Similarly, although Weiner (1999, 2000) recently argued that Rorschach scores are generally unrelated to diagnoses, it was only a few years ago that Weiner (1997) stated

At present the Rorschach Comprehensive System provides indices for schizophrenia (*SCZI*) and depression (*DEPI*) that can prove helpful in identifying these two conditions. . . . Recent work by Gacono and Meloy (1994) suggested that a similarly sound and useful index of psychopathic personality can now be constructed. . . . In addition, although further documentation is needed, accumulating data indicate that there are on the horizon adequately conceptualized and empirically valid Rorschach indices for bipolar disorder, borderline and schizotypal personality disorder, and acute and chronic stress disorder. (pp. 10–11)

Even in this issue of the *Journal of Personality Assessment*, both Ganellen and Gacono et al. adopt the position that certain Rorschach scores are related to diagnostic categories such as depression, ASPD, CD, and psychopathy. Thus, our critics are clearly at odds with each other concerning the fundamental meaning of Rorschach findings. Ganellen and Gacono et al. maintain that the Rorschach should be moderately or highly correlated with measures of psychiatric diagnoses, whereas Bornstein (this issue) maintains that the Rorschach should be uncorrelated with these measures. This state of affairs is confusing at best.

Not only Ganellen (this issue) and Gacono et al. (this issue), but many other researchers and clinicians have assumed that numerous Rorschach scores are related to psychiatric diagnoses. As our review on this topic showed (Wood et al., 2000), hundreds of studies have examined the relationships between the Rorschach scores and psychiatric diagnoses. Virtually all of these researchers hypothesized that they would find positive relations, although many or most came up empty-handed.

Only now, when the research has yielded overwhelmingly negative results, have Rorschach proponents such as Bornstein (this issue) and Weiner (1999) taken the position that the test really should not be related to diagnoses after all. Sir Karl Popper (1974), the noted philosopher of science, argued that such post hoc at-

tempts to minimize or explain away negative research findings rarely, if ever, constitute good science:

We can always adopt evasive tactics in the face of refutations. ... We can always immunize a theory against refutation. There are many such evasive immunizing tactics; and if nothing better occurs to us, we can always deny the objectivity—or even the existence—of the refuting observation. (p. 983)

Popper (1974) did not adopt an inflexible rule in this respect and even suggested that it is sometimes good science to defend preconceived ideas in the face of apparently disconfirming evidence. Meehl (1990), however, argued that this strategy is justified only when a research program or theory possesses an excellent track record of previously corroborated predictions (i.e., “money in the bank;” see also Lakatos, 1970). It is difficult to maintain that the Rorschach possesses such a track record. Moreover, we believe that in the case of the Rorschach and diagnoses, the negative results are so strong and numerous that Bornstein and other Rorschach proponents should be questioning the validity of at least some of the measure’s indexes and scores, rather than minimizing the importance of the research findings.

The (Lack of) Relationship Between the Rorschach and Self-Report Measures

Bornstein (this issue) does not question the Rorschach’s validity, despite its failure to demonstrate a positive relation with most psychiatric diagnoses. Furthermore, he argues that the test’s failure to demonstrate a relation with self-report questionnaires and interviews actually constitutes a strength:

In the case of the RIM, modest correlations with scores on questionnaire or interview measures have occasionally been presented as evidence against the convergent validity of the test (see Wood et al., 1996). In fact, the opposite is true: Such findings support the discriminant validity of the measure. (p. 45)

In contrast to the commonly accepted view that a purported measure of a construct should correlate with other measures of the same construct (Campbell & Fiske, 1959), Bornstein argues that just the opposite is true: A failure to correlate with other measures constitutes positive evidence of (discriminant) validity.

To see the difficulty in Bornstein’s (this issue) argument, it is helpful to turn to another of his articles, which is on the construct validity of the Rorschach Oral Dependency scale (*ROD*; Bornstein, 1996). In that article, Bornstein (1996, p. 202) noted that correlations between the *ROD* and a self-report measure of dependency

ranged from .32 to .67 in women, and from .37 to .48 in men. Bornstein (1996) cited these positive correlations as evidence of “convergent validity” for the *ROD*.

Although Bornstein (1996) held forth positive correlations between the Rorschach and self-report measures as evidence of convergent validity, Bornstein (this issue) also argues that “modest” correlations are evidence of “discriminant validity.” Of course, such a position allows virtually any correlational finding to be interpreted as support for Rorschach validity. A positive correlation can be interpreted as evidence of convergent validity, whereas a modest or null correlation can be interpreted as evidence of discriminant validity. Again, we suggest that such an approach is contrary to the principles of sound scientific reasoning. If opposite results from the same study can both be interpreted as support for the Rorschach, then it is impossible for the study to “falsify” (Popper, 1974) hypotheses concerning the validity of the Rorschach. As Popper noted, a theory that is consistent with every conceivable outcome is untestable and therefore unscientific.

Conclusions: The Validity and Utility of Negative Findings Regarding the Rorschach

Bornstein (this issue) argues that “efforts to evaluate the validity and utility of the RIM by assessing the degree to which RIM scores correlate with *DSM-IV* diagnoses and symptom ratings are misguided and inappropriate” (p. 44). In conclusion, we express our strong disagreement with Bornstein on this issue. We believe that the numerous negative findings regarding the Rorschach and psychiatric diagnoses have a great deal to tell psychologists about the validity and utility of the Rorschach. First, the negative research findings have clarified a fact that was not obvious 5 years ago, even to most Rorschach experts: With only a few exceptions, the Rorschach (including the CS) has little usefulness for detecting psychiatric diagnoses. This important new piece of knowledge has substantial clinical implications, for example, in managed care and forensic settings.

Second, the negative findings are informative regarding the construct validity of CS scores. For example, inanimate movement responses (*m*), diffuse shading responses (*Y*), and the *D* score are supposedly related to anxiety and stress (McCain, Fink, Gallina, & Johnson, 1992; Perry et al., 1995; but see Frank, 1978, 1993a, 1993b). Yet research does not indicate that these Rorschach variables are significantly related to PTSD or other anxiety disorders (Wood et al., 2000). To our thinking, these null findings strongly suggest that *m*, *Y*, and *D* are not related to anxiety and stress at all. Perhaps Bornstein (this issue) or other Rorschach proponents would respond to these negative findings by arguing that the Rorschach variables in fact assess anxiety and stress, but not any *DSM-IV* diagnoses that are associated with anxiety and stress. If so, the burden of proof falls squarely on these proponents to demonstrate that the Rorschach variables correlate consistently with other well-established anxiety measures.

Similarly, Space (*S*) responses are supposedly related to negativism, oppositionality, and an angry attitude toward the environment (Exner, 1991, p. 199), whereas an absence of Texture (*T*) responses is thought to indicate reluctance “about creating or maintaining close emotional ties with others” (Exner, 1991, p. 184). One would therefore expect Space responses and Textureless protocols to be more common among psychopaths or individuals with ASPD or CD than among other individuals (Gacono & Meloy, 1994; Weber et al., 1992). However, as we discussed earlier in this article, research has not corroborated these hypotheses. Again, the null findings seem to indicate that *S* and *T* do not measure what they are purported to.

A great deal of research has examined the relationship of the Rorschach to psychiatric diagnoses and self-report questionnaires. Contrary to Bornstein (this issue), we do not believe that the predominantly negative findings are irrelevant or should be interpreted as providing support for the test’s validity. To the contrary, these findings constitute important disconfirming scientific evidence and must be weighed carefully in any evaluation of the test’s validity.

ACKNOWLEDGMENT

We thank Robert Archer, Radhika Krishnamurthy, John Hunsley, and Michael Bailey for sharing valuable information with us that was incorporated into our article.

REFERENCES

- Abraham, P. P., Lepisto, B. L., Lewis, M. G., Schultz, L., & Finkelberg, S. (1994). An outcome study: Changes in Rorschach variables of adolescents in residential treatment. *Journal of Personality Assessment*, *62*, 505–514.
- Acklin, M. W., McDowell, C. J., Verschell, M. S., & Chan, D. (2000). Interobserver agreement, intraobserver reliability, and the Rorschach Comprehensive System. *Journal of Personality Assessment*, *74*, 15–47.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Anderson, T., & Dixon, W. E. (1993). The factor structure of the Rorschach for adolescent inpatients. *Journal of Personality Assessment*, *60*, 319–332.
- Archer, R. P. (Ed.). (1999). Perspectives on the Rorschach [Special section]. *Assessment*, *6*(3).
- Archer, R. P., & Gordon, R. A. (1988). MMPI and Rorschach indices of schizophrenic and depressive diagnoses among adolescent inpatients. *Journal of Personality Assessment*, *52*, 276–287.
- Archer, R. P., & Krishnamurthy, R. (1993a). Combining the Rorschach and the MMPI in the assessment of adolescents. *Journal of Personality Assessment*, *60*, 132–140.
- Archer, R. P., & Krishnamurthy, R. (1993b). A review of MMPI and Rorschach interrelationships in adult samples. *Journal of Personality Assessment*, *61*, 277–293.
- Archer, R. P., & Krishnamurthy, R. (1997). MMPI–A and Rorschach indices related to depression and conduct disorder: An evaluation of the incremental validity hypothesis. *Journal of Personality Assessment*, *69*, 517–533.

- Ball, J. D., Archer, R. P., Gordon, R. A., & French, J. (1991). Rorschach depression indices with children and adolescents: Concurrent validity findings. *Journal of Personality Assessment*, *57*, 465–476.
- Barlow, D. H. (Ed.). (2001). *Clinical Psychology: Science and Practice*, *8*(3).
- Beutler, L. E. (Ed.). (2000). *Journal of Clinical Psychology*, *56*(3).
- Bornstein, R. F. (1996). Construct validity of the Rorschach Oral Dependency scale: 1967–1995. *Psychological Assessment*, *8*, 200–205.
- Brinkman, D. C., Overholser, J. C., & Klier, D. (1994). Emotional distress in adolescent psychiatric inpatients: Direct and indirect measures. *Journal of Personality Assessment*, *62*, 472–484.
- Butcher, J. N., Williams, C. L., Graham, J. R., Archer, R. P., Tellegen, A., Ben-Porath, Y. S., & Kaemmer, B. (1992). *MMPI-A (Minnesota Multiphasic Personality Inventory-Adolescent): Manual for administration, scoring, and interpretation*. Minneapolis: University of Minnesota Press.
- Caine, S. L., Frueh, B. C., & Kinder, B. N. (1995). Rorschach susceptibility to malingered depressive disorders in adult females. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment*. (Vol. 10, pp. 165–174). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105.
- Carter, C. L., & Dacey, C. M. (1996). Validity of the Beck Depression Inventory, MMPI, and Rorschach in assessing adolescent depression. *Journal of Adolescence*, *19*, 223–231.
- Cotugno, A. J. (1995). Personality attributes of Attention Deficit Hyperactivity Disorder (ADHD) using the Rorschach Inkblot Test. *Journal of Clinical Psychology*, *51*, 554–561.
- Cronbach, L. J. (1960). *Essentials of psychological testing* (3rd ed.). New York: Harper.
- Darcangelo, S. M. (1997). Psychological and personality correlates of the Massachusetts treatment center classification system for rapists (Doctoral dissertation, Simon Fraser University, 1996). *Dissertation Abstracts International*, *58*, 2115B.
- Egozi-Profeta, V. L. (1999). A comparison of the Roemer and the Rorschach tests as tools for distinguishing characteristics of psychopathy (Doctoral dissertation, Miami Institute of Psychology of the Caribbean Center for Advanced Studies, 1998). *Dissertation Abstracts International*, *60*, 1345B.
- Exner, J. E., Jr. (1986). *The Rorschach: A comprehensive system: Vol. 1. Basic foundations* (2nd ed.). New York: Wiley.
- Exner, J. E., Jr. (1991). *The Rorschach: A comprehensive system: Vol. 2. Interpretation* (2nd ed.). New York: Wiley.
- Exner, J. E., Jr. (1993). *The Rorschach: A comprehensive system: Vol. 1. Basic foundations* (3rd ed.). New York: Wiley.
- Exner, J. E., Jr., & Weiner, I. B. (1995). *The Rorschach: A comprehensive system: Vol. 3. Assessment of children and adolescents* (2nd ed.). New York: Wiley.
- Frank, G. (1978). On the validity of hypotheses derived from the Rorschach: III. The relationship between shading and anxiety. *Perceptual and Motor Skills*, *46*, 531–538.
- Frank, G. (1993a). On the meaning of movement responses on the Rorschach. *Psychological Reports*, *73*, 1219–1225.
- Frank, G. (1993b). On the validity of hypotheses derived from the Rorschach: The relationship between shading and anxiety, update 1992. *Psychological Reports*, *72*, 519–522.
- Gacono, C. B. (1990). An empirical study of object relations and defensive operations in antisocial personality disorder. *Journal of Personality Assessment*, *54*, 589–600.
- Gacono, C. B., & Meloy, J. R. (1991). A Rorschach investigation of attachment and anxiety in antisocial personality disorder. *Journal of Nervous and Mental Disease*, *179*, 546–552.
- Gacono, C. B., & Meloy, J. R. (1992). The Rorschach and the DSM-III-R antisocial personality: A tribute to Robert Lindner. *Journal of Clinical Psychology*, *48*, 393–406.
- Gacono, C. B., & Meloy, J. R. (1994). *The Rorschach assessment of aggressive and psychopathic personalities*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

- Gacono, C. B., Meloy, J. R., & Berg, J. L. (1992). Object relations, defensive operations, and affective states in narcissistic, borderline, and antisocial personality disorder. *Journal of Personality Assessment*, *59*, 32–49.
- Gacono, C. B., Meloy, J. R., & Heaven, T. R. (1990). A Rorschach investigation of narcissism and hysteria in antisocial personality. *Journal of Personality Assessment*, *55*, 270–279.
- Garb, H. N., Florio, C. M., & Grove, W. M. (1998). The validity of the Rorschach and the Minnesota Multiphasic Personality Inventory: Results from meta-analyses. *Psychological Science*, *9*, 402–404.
- Garb, H. N., Wood, J. M., Nezworski, M. T., Grove, W. M., & Stejskal, W. J. (in press). Towards a resolution of the Rorschach controversy. *Psychological Assessment*.
- Greenwald, D. F. (1990). An external construct validity study of Rorschach personality variables. *Journal of Personality Assessment*, *55*, 768–780.
- Greenwald, D. F. (1991). Personality dimensions reflected by the Rorschach and the 16PF. *Journal of Clinical Psychology*, *47*, 708–715.
- Greenwald, D. F. (1999). Relationships between the Rorschach and the NEO-Five Factor Inventory. *Psychological Reports*, *85*, 519–527.
- Gronnerod, C. (1999). Rorschach interrater agreement estimates: An empirical evaluation. *Scandinavian Journal of Psychology*, *40*, 115–120.
- Groth-Marnat, G. (1997). *Handbook of psychological assessment* (3rd ed.). New York: Wiley.
- Hamel, M., Shaffer, T. W., & Erdberg, P. (2000). A study of nonpatient preadolescent Rorschach protocols. *Journal of Personality Assessment*, *75*, 280–294.
- Hare, R. (1980). A research scale for the assessment of psychopathy in criminal populations. *Personality and Individual Differences*, *1*, 111–119.
- Hare, R. (1991). *Manual for the Revised Psychopathy Checklist*. Toronto, Canada: Multihealth Systems.
- Harris, G. T., Rice, M. E., & Quinsey, V. L. (1994). Psychopathy as a taxon: Evidence that psychopaths are a discrete class. *Journal of Consulting and Clinical Psychology*, *62*, 387–397.
- Hiller, J. B., Rosenthal, R., Bornstein, R. F., Berry, D. T. R., & Brunell-Neuleib, S. (1999). A comparative meta-analysis of Rorschach and MMPI validity. *Psychological Assessment*, *11*, 278–296.
- Hunsley, J., & Bailey, J. M. (in press). Whither the Rorschach? An analysis of the evidence. *Psychological Assessment*.
- Ilonen, T., Taiminen, T., Karlsson, H., Lauerma, H., Leinonen, K.-M., Wallenius, E., Tuimala, P., & Salokangas, R. (1999). Diagnostic efficiency of the Rorschach schizophrenia and depression indices in identifying first-episode schizophrenia and severe depression. *Psychiatry Research*, *87*, 183–192.
- Jansak, D. M. (1997). The Rorschach Comprehensive System Depression Index, depression heterogeneity, and the role of self-schema (Doctoral dissertation, California School of Professional Psychology, 1996). *Dissertation Abstracts International*, *57*, 6576B.
- Jorgensen, K., Andersen, T. J., & Dam, H. (2000). The diagnostic efficiency of the Rorschach Depression Index and the Schizophrenia Index: A review. *Assessment*, *7*, 259–280.
- Karfigin, L. D. (1989). Conduct disorder and depression in adolescence (Doctoral dissertation, University of Maryland, Baltimore County, 1988). *Dissertation Abstracts International*, *49*, 4575B.
- Kessel, J. B., & Zimmerman, M. (1993). Reporting errors in studies of the diagnostic performance of self-administered questionnaires: Extent of the problem, recommendations for standardized presentation of results, and implications for the peer review process. *Psychological Assessment*, *5*, 395–399.
- Krishnamurthy, R., & Archer, R. P. (1997, March). An incremental validity evaluation of the MMPI-A and Rorschach. In B. L. Smith (Chair), *Controversial discussion: Rorschach and MMPI convergence/divergence*. Symposium presented at the 1997 meeting of the Society for Personality Assessment, San Diego, CA.

- Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–95). Cambridge, England: Cambridge University Press.
- Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology*, *68*, 151–158.
- Lilienfeld, S. O. (1994). Conceptual problems in the assessment of psychopathy. *Clinical Psychology Review*, *14*, 17–38.
- Lilienfeld, S. O. (1998). Methodological advances and developments in the assessment of psychopathy. *Behaviour Research and Therapy*, *36*, 99–125.
- Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality in noncriminal populations. *Journal of Personality Assessment*, *66*, 488–524.
- Lipovsky, J. A., Finch, A. J., & Belter, R. W. (1989). Assessment of depression in adolescents: Objective and projective measures. *Journal of Personality Assessment*, *53*, 449–458.
- Long, J. S. (1995). A comparative study between the Rorschach and MMPI–A in predicting symptoms and behaviors with inpatient adolescents (Doctoral dissertation, University of Tulsa, 1995). *Dissertation Abstracts International*, *56*, 3452B.
- Loving, J. L. (1998). Selected Rorschach variables of psychopathic male juvenile offenders (Doctoral dissertation, Widener University, 1998). *Dissertation Abstracts International*, *59*, 0878B.
- Loving, J. L., & Russell, W. F. (2000). Selected Rorschach variables of psychopathic juvenile offenders. *Journal of Personality Assessment*, *75*, 126–142.
- McCain, W., Fink, A. D., Gallina, H., & Johnson, J. (1992). Effects of laboratory-induced controllable and uncontrollable stress on Rorschach variables *m* and *Y*. *Journal of Personality Assessment*, *59*, 564–573.
- Meehl, P. E. (1990). Appraising and amending theories: The strategy of Lakatosian defense and two principles that warrant it. *Psychological Inquiry*, *1*, 108–141.
- Meehl, P. E., & Golden, R. R. (1982). Taxometric methods. In P. C. Kendall & J. N. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 127–181). New York: Guilford.
- Meehl, P. E., & Rosen, A. (1955). Antecedent probability of psychiatric signs, patterns, or cutting scores. *Psychological Bulletin*, *52*, 194–216.
- Meloy, J. R., & Gacono, C. B. (1995). Assessing the psychopathic personality. In J. N. Butcher (Ed.), *Clinical personality assessment* (pp.410–422). New York: Oxford University Press.
- Meloy, J. R., & Gacono, C. B. (2000). Assessing psychopathy: Psychological testing and report writing. In C. B. Gacono (Ed.), *The clinical and forensic assessment of psychopathy: A practitioner's guide* (pp. 231–249). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Meyer, G. J. (1991). An empirical search for fundamental personality and mood dimensions within the Rorschach test (Doctoral dissertation, Loyola University of Chicago, 1989). *Dissertation Abstracts International*, *52*, 1071B–1072B.
- Meyer, G. J. (1992). The Rorschach's factor structure: A contemporary investigation and historical review. *Journal of Personality Assessment*, *59*, 117–136.
- Meyer, G. J. (1993). The impact of response frequency on the Rorschach constellation indices and on their validity with diagnostic and MMPI–2 criteria. *Journal of Personality Assessment*, *60*, 153–180.
- Meyer, G. J. (1996). The Rorschach and MMPI: Toward a more scientifically differentiated understanding of cross-method assessment. *Journal of Personality Assessment*, *67*, 558–578.
- Meyer, G. J. (1997a). Assessing reliability: Critical corrections for a critical examination of the Rorschach Comprehensive System. *Psychological Assessment*, *9*, 480–489.
- Meyer, G. J. (1997b). Thinking clearly about reliability: More critical corrections regarding the Rorschach Comprehensive System. *Psychological Assessment*, *9*, 495–498.
- Meyer, G. J. (Ed.). (1999). Utility of the Rorschach for clinical assessment [Special section]. *Psychological Assessment*, *11*(3).
- Muntz, A. (1999). Object relations and defense mechanisms of psychopathic and nonpsychopathic female offenders: A descriptive study (Doctoral dissertation, California School of Professional Psychology, 1998). *Dissertation Abstracts International*, *60*, 2954B.

- Murphy-Peaslee, D. M. (1995). An investigation of incarcerated females: Rorschach indices and Psychopathy Checklist scores (Doctoral dissertation, California School of Professional Psychology, 1993). *Dissertation Abstracts International*, 56, 0531B.
- Myers, B. P. (Ed.). (2001). *Journal of Forensic Psychology Practice*, 1(3).
- Nakata, L. M. (2000). Interrater reliability and the Comprehensive System for the Rorschach: Clinical and non-clinical protocols (Doctoral dissertation, Pacific Graduate School of Psychology, 1999). *Dissertation Abstracts International*, 60, 4296B.
- Nezworski, M. T., & Wood, J. M. (1995). Narcissism in the Comprehensive System for the Rorschach. *Clinical Psychology: Science and Practice*, 2, 179–199.
- Parker, K. C. H., Hanson, R. K., & Hunsley, J. (1988). MMPI, Rorschach, and WAIS: A meta-analytic comparison of reliability, stability, and validity. *Psychological Bulletin*, 103, 367–373.
- Perry, W., Sprock, J., Schaible, D., McDougall, A., Minassian, A., Jenkins, M., & Braff, D. (1995). Amphetamine on Rorschach measures in normal subjects. *Journal of Personality Assessment*, 64, 456–465.
- Ponder, J. I. (1999). An investigation of psychopathy in a sample of violent juvenile offenders (Doctoral dissertation, University of Texas at Austin, 1998). *Dissertation Abstracts International*, 59, 5105B.
- Popper, K. (1974). Replies to my critics. In P. A. Schilpp (Ed.), *The philosophy of Karl Popper* (pp. 961–1197). La Salle, IL: Open Court.
- Psychological Corporation. (1997). *Wechsler Adult Intelligence Scale, (3rd ed.)*. *Wechsler Memory Scale, (3rd ed.)*. *Technical manual*. San Antonio, TX: Author.
- Ritzler, B. (1995). Putting your eggs in the content analysis basket: A response to Aronow, Reznikov, and Moreland. *Journal of Personality Assessment*, 64, 229–234.
- Ryan, J. J., Prifitera, A., & Powers, L. (1983). Scoring reliability on the WAIS–R. *Journal of Consulting and Clinical Psychology*, 51, 149–150.
- Sells, J. E. (1991). A validity study of the *DEPI* index: The Rorschach Comprehensive System (Doctoral dissertation, University of Utah, 1990). *Dissertation Abstracts International*, 51, 5590B.
- Shaffer, T. W., Erdberg, P., & Haroian, J. (1999). Current nonpatient data for the Rorschach, WAIS–R, and MMPI–2. *Journal of Personality Assessment*, 73, 305–316.
- Siemsen, R. A. (1999). Relationships of Rorschach and MMPI–2 variables to the Hare Psychopathy Checklist–Revised among mentally-ill incarcerated felons (Doctoral dissertation, California School of Professional Psychology, 1999). *Dissertation Abstracts International*, 60, 2367B.
- Silberg, J. L., & Armstrong, J. G. (1992). The Rorschach test for predicting suicide among depressed adolescent inpatients. *Journal of Personality Assessment*, 59, 290–303.
- Singer, H. K., & Brabender, V. (1993). The use of the Rorschach to differentiate unipolar and bipolar disorders. *Journal of Personality Assessment*, 60, 333–345.
- Slate, J. R., Jones, C. R., Coulter, C., & Covert, T. L. (1992). Practitioners' administration and scoring of the WISC–R: Evidence that we do err. *Journal of School Psychology*, 30, 77–82.
- Slate, J. R., Jones, C. H., Murray, R. A., & Coulter, C. (1993). Evidence that practitioners err in administering and scoring the WAIS–R. *Measurement and Evaluation in Counseling and Development*, 25, 156–161.
- Smith, A. M. (1995). Juvenile psychopathy: Rorschach assessment of narcissistic traits in conduct disordered adolescents (Doctoral dissertation, California School of Professional Psychology, 1994). *Dissertation Abstracts International*, 55, 5088B.
- Smith, A. M., Gacono, C. B., & Kaufman, L. (1997). A Rorschach comparison of nonpsychopathic conduct disordered adolescents. *Journal of Clinical Psychology*, 53, 289–300.
- Smith, A. M., Gacono, C. B., & Kaufman, L. (1998). Erratum: A Rorschach comparison of psychopathic and nonpsychopathic conduct disordered adolescents. *Journal of Clinical Psychology*, 54, 1151.
- Viglione, D. J., Brager, R. C., & Haller, N. (1988). Usefulness of structural Rorschach data in identifying inpatients with depressive symptoms: A preliminary study. *Journal of Personality Assessment*, 52, 524–529.
- Viglione, D. J., Brager, R., & Haller, N. (1991). Psychoanalytic interpretation of the Rorschach: Do we have better hieroglyphics? *Journal of Personality Assessment*, 57, 1–9.
- Weber, C. A., Meloy, J. R., & Gacono, C. B. (1992). A Rorschach study of attachment and anxiety in inpatient conduct-disordered and dysthymic adolescents. *Journal of Personality Assessment*, 58, 16–26.

- Weiner, I. B. (1997). Current status of the Rorschach Inkblot Method. *Journal of Personality Assessment*, *68*, 5–19.
- Weiner, I. B. (1998). *Principles of Rorschach interpretation*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Weiner, I. B. (1999). What the Rorschach can do for you: Incremental validity in clinical applications. *Assessment*, *6*, 327–338.
- Weiner, I. B. (2000). Using the Rorschach properly in practice and research. *Journal of Clinical Psychology*, *56*, 435–438.
- Welsh, R. K. (1999). Psychopathy and psychological risk markers of violent recidivism (Doctoral dissertation, Rosemead School of Psychology, Biola University, 1999). *Dissertation Abstracts International*, *60*, 2968B.
- Whitten, J., Slate, J. R., Jones, C. H., & Shine, A. E. (1994). Examiner errors in administering and scoring the WPPSI-R. *Journal of Psychoeducational Assessment*, *12*, 49–54.
- Wood, J. M., & Lilienfeld, S. O. (1999). The Rorschach Inkblot Test: A case of overstatement? *Assessment*, *6*, 341–349.
- Wood, J. M., Lilienfeld, S. O., Garb, H. N., & Nezworski, M. T. (2000). The Rorschach Test in clinical diagnosis: A critical review, with a backward look at Garfield (1947). *Journal of Clinical Psychology*, *56*, 395–430.
- Wood, J. M., Nezworski, M. T., Garb, H. N., & Lilienfeld, S. O. (in press). The misperception of psychopathology: Problems with the norms of the Comprehensive System for the Rorschach. *Clinical Psychology: Science and Practice*.
- Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1996). The Comprehensive System for the Rorschach: A critical examination. *Psychological Science*, *7*, 3–10.
- Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1997). The reliability of the Comprehensive System for the Rorschach: A Comment on Meyer (1997). *Psychological Assessment*, *9*, 490–494.
- Wood, J. M., Nezworski, M. T., Stejskal, W. J., Garven, S., & West, S. G. (1999). Methodological issues in evaluating Rorschach validity: A comment on Burns and Viglione (1996), Weiner (1996), and Ganellen (1996). *Assessment*, *6*, 115–120.
- Young, M. H., Justice, J. V., Erdberg, P. S., & Gacono, C. B. (2000). The incarcerated psychopath in psychiatric treatment: Management or treatment. In C. B. Gacono (Ed.), *The clinical and forensic assessment of psychopathy: A practitioner's guide* (pp. 313–331). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Zagon, I. K., & Jackson, H. J. (1994). Construct validity of a psychopathy measure. *Personality and Individual Differences*, *17*, 125–135.

James M. Wood
 Department of Psychology
 University of Texas
 El Paso TX 79968
 E-mail: jawood@utep.edu

Received November 6, 2000