The Misperception of Psychopathology: Problems With the Norms of the Comprehensive System for the Rorschach

James M. Wood, University of Texas at El Paso
M. Teresa Nezworski, University of Texas at Dallas
Howard N. Garb, Pittsburgh V.A. Health Care System and University of Pittsburgh
Scott O. Lilienfeld, Emory University

The norms of the Comprehensive System (CS) for the Rorschach have been praised for their scientific and clinical value. However, recent findings by Shaffer, Erdberg, and Haroian suggest that the CS norms are inaccurate. We examined 14 CS variables in 32 studies of nonpatient adults. The participants in these studies exhibited statistically and clinically significant differences from CS nonpatient norms for all 14 variables. Overall, nonpatient adults tended to appear pathological when compared with the CS norms. The implications of these findings for misdiagnosis and the misperception of psychopathology are discussed. Use of the CS norms in clinical or forensic settings is inadvisable. Commonly held beliefs about the validity of CS variables are likely to be false and in need of revision.

Key words: **Rorschach, Comprehensive System, projective tests, norms, overpathologizing.** ***[Clin Psychol Sci Prac 8:350–373, 2001]***

As Hunsley and Bailey (1999) have noted, the Rorschach Inkblot test "has the dubious distinction of being, simultaneously, the most cherished and the most reviled of all psychological assessment instruments" (p. 266). In the 1950s and 1960s, the test aroused heated controversy among psychologists. Critics argued that the Rorschach lacked standardized administration procedures and adequate norms and that evidence for reliability and validity was often weak or nonexistent (Eysenck, 1959; Jensen, 1965; Shaffer, 1959). Defenders responded that clinical experience had confirmed the test's value and that many criticisms of it were "naive and unjust, often fomented from bias, ignorance, or simply a misunderstanding of the method and the principles that led to its exploration by Rorschach" (Exner, 1993, p. 3).

In the absence of firm scientific support, and despite the enthusiasm of many clinicians, the Rorschach might have sunk slowly into oblivion. However, Exner's 1974 publication of *The Rorschach: A Comprehensive System* (*TRACS*) dramatically revived the test's psychometric status. This book, along with its subsequent extensions and revisions (Exner, 1978, 1986, 1991, 1993; Exner & Weiner, 1982, 1995), seemed at last to establish the Rorschach on a firm scientific foundation. Exner's Comprehensive System (CS) for the Rorschach provided detailed rules for administration and scoring and an impressive set of norms for both children and adults. Furthermore, various editions of *TRACS* reported strikingly positive findings from hundreds of unpublished reliability and validity studies by Exner's Rorschach Workshops.

The achievements of the CS elicited widespread praise from the psychological community. Anastasi (1988) declared that "The availability of this system, together with the research completed thus far, has injected new life into the Rorschach as a potential psychometric instrument" (p. 599). The Board of Professional Affairs (1998, p. 392) of the American Psychological Association commended Exner for his "resurrection" of the test. In the 1990s, surveys indicated that the Rorschach was widely used in clinical and forensic settings and that the CS was the most widely used scoring system for the Rorschach (Ackerman & Ackerman, 1997; Lees-Haley, 1992; Pinkerman, Haynes & Keiser, 1993; Piotrowski, 1999; Watkins, Campbell, Nieberding, & Hallmark, 1995).

Then, in the mid-1990s, the Rorschach controversy unexpectedly revived. For the first time in its history the CS became the subject of vigorous debate. Critics pointed

Address correspondence to James M. Wood, Ph.D., Department of Psychology, University of Texas at El Paso, El Paso, TX 79968. Electronic mail may be sent to jawood@utep.edu.

out that the unpublished reliability and validity studies of Exner's Rorschach Workshops were often unavailable for scrutiny by independent scholars, and that many CS scores lacked well-demonstrated validity (Nezworski & Wood, 1995; Wood, Nezworski, & Stejskal, 1996a, 1996b; but see Exner, 1995a, 1996). Additional debates flared between Rorschach advocates and critics regarding such fundamental issues as scoring reliability, test–retest reliability, incremental validity, clinical utility, effects of method variance, cultural sensitivity, and research methodology (Acklin, 1999; Archer, 1999; Aronow, 1999; Aronow, Reznikoff, & Moreland, 1994, 1995; Costello, 1999; Dawes, 1994; Ganellen, 1996a, 1996b; Gann, 1995; Garb, 1998, 1999; Garb, Wood, & Nezworski, 2000; Garb, Wood, Nezworski, Grove, & Stejskal, in press; Garfield, 2000; Hunsley & Bailey, 1999; Jorgensen, Andersen, & Dam, in press; Kubiszyn et al., 2000; Lerner, 2000; Lilienfeld, Wood, & Garb, 2000; Sechrest & McKnight, 2000; Sechrest, Stickle, & Stewart, 1998; Stricker & Gold, 1999; Viglione, 1999; Weiner, 1996, 1999, 2000; Wood & Lilienfeld, 1999; Wood, Lilienfeld, Garb, & Nezworski, 2000a, 2000b; Wood, Nezworski, Stejskal, Garven, & West, 1999).

The controversy over the Rorschach CS has garnered an increasing amount of attention: In 1999 and 2000 three clinical journals (*Assessment, Psychological Assessment, Journal of Clinical Psychology*) published debates between critics and proponents. The widespread scientific acceptance of the CS that characterized the 1970s and 1980s has vanished, and one scholar has called for a moratorium on use of the Rorschach, including the CS, in clinical and forensic settings (Garb, 1999). As Robert Archer (1999), editor of *Assessment,* commented, "the assumption that the Rorschach Comprehensive System rests solidly and uniformly on an empirical foundation has been forced to undergo a significant re-examination" (p. 309).

In the midst of these heated exchanges, one point has gone unchallenged. Both critics (e.g., Wood & Lilienfeld, 1999) and proponents (e.g., Weiner, 1998) of the Rorschach have usually assumed that the norms of the CS, published over a period of 20 years, represent a substantial scientific and clinical achievement. Although problematic aspects of the CS norms have sometimes been identified (Loucks, Burstein, Boros, & Kregor, 1980; Vincent & Harman, 1991; Wood & Lilienfeld, 1999), such issues have attracted little attention in the recent debate regarding the CS.

Thus, both proponents and critics of the Rorschach were caught by surprise in late 1999 when Shaffer, Erdberg, and Haroian (1999) presented data that seemed to call the CS norms into question. In a sample of 123 nonpatient adults from California, Shaffer and his colleagues found substantial discrepancies from the CS norms for a large number of important Rorschach variables. For example, about one in six of the Shaffer et al. nonpatient subjects scored in the pathological range (>4) on the Schizophrenia Index (*SCZI*). More than one-fourth of their nonpatients (29%) gave at least one Reflection response, a supposedly rare Rorschach indicator of narcissism (Exner, 1991, p. 149). Contrary to what would be expected, the nonpatients appeared seriously disturbed on Rorschach measures of perceptual accuracy and distorted thinking. For example, the mean score of Conventional Form (*X+%*) reported for the Shaffer et al. sample was more than three standard deviations below the CS norms (.51 versus.79), whereas the mean score of Distorted Form (*X–%*) was more than two standard deviations higher than the norms (.21 versus.07). Substantial discrepancies were reported for other Rorschach indicators of emotional functioning or psychopathology, including EB style (percentage of ambitent protocols), the Affective Ratio (*Afr*), the Form-Color Ratio (*FC: CF+C*), Popular Responses, Diffuse Shading Responses (*Sum Y*), Texture responses ($T = 0$), the Weighted Sum of Color Responses (*WSumC*), Morbid responses (MOR), the Weighted Sum of the 6 Special Scores (*WSum6*), Lambda >.99, and Pure Human responses (*Pure H* < 2). Although these discrepancies were sizable and probably statistically significant, Shaffer et al. did not report statistical tests. Nearly all the discrepancies had the effect of making the nonpatient group appear maladjusted compared with the normative data.

The article by Shaffer et al. (1999) provided the impetus for the present review. We set out to examine whether the results of researchers other than Shaffer et al. suggest similar discrepancies from the CS norms. We identified 32 studies in which researchers had collected CS scores from nonpatient American adults. The scores for 14 Rorschach variables were extracted from these studies, combined, and then compared with current CS nonpatient norms (Exner, 1993). We hypothesized that the Rorschach scores of the aggregated nonpatient adults would differ significantly from the CS norms in a manner consistent with the findings of Shaffer et al., and that the nonpatients would generally appear pathological compared with these norms.

COLLECTION OF STUDIES FOR THE REVIEW

Search for Studies

We conducted a search for all CS studies that have examined samples of nonpatient adults. Samples were excluded if (a) the CS was not used for both administration and scoring of protocols, (b) the sample consisted of psychiatric or psychotherapy patients, or individuals with current psychiatric diagnoses; (c) the group consisted of prisoners, delinquents, psychopaths, convicted or accused criminals, or civil litigants; (d) the group consisted of individuals who had admitted or were suspected of sexual misconduct (e.g., sex abusers) or substance abuse; or (e) the group consisted of deaf or color-blind participants. The aim of the present study was to make comparisons with the CS norms (Exner, 1993), which are based on protocols sampled from the Rorschach Workshops subject pool. To ensure that the subjects in the present study were entirely separate from the CS normative sample, all samples that potentially contained subjects from the Rorschach Workshops subject pool were excluded. Specifically, samples were excluded if (f) they had been sampled entirely or in part from the subjects of the Rorschach Workshops, or from studies by Exner.

The search consisted of three steps. First, a manual search was conducted of all articles published in the *Journal of Personality Assessment* from 1974 through June 1999 to identify all studies that had included nonpatient adult participants. In this step, eight studies were identified that met criteria for inclusion in the present study. Second, a search was made of the PsycINFO database for all articles or book chapters whose title or abstract contained the word "Rorschach" for the years 1974 through 1998 (1974 was chosen because the CS was introduced in that year). The abstracts of all articles were read, and studies were identified that included nonpatient participants. In this step, an additional five articles and two book chapters were identified that met criteria for inclusion.

Third, we manually searched 83 Rorschach dissertations for nonpatient adult samples of subjects. These dissertations had been accumulated by J.M.W. over the prior 6 years for several literature reviews on topics related to the Rorschach, including (a) depression, (b) narcissism, (c) cultural diversity, (d) posttraumatic stress disorder, (e) psychopathy and criminal behavior, (f) child abuse, and (g) the Human Experience variable (Garb, Wood, & Lilienfeld, 2000; Garb et al., in press; Lilienfeld et al., 2000; Nezworski & Wood, 1995; Wood & Lilienfeld, 1999; Wood et al., 1999, 2000a). These dissertations did not constitute a randomly selected sample. However, they had originally been obtained for reasons unrelated to the purposes of the present study, and there was no reason to expect that the means and standard deviations of the nonpatient samples included in them would be systematically different from the means and standard deviations of nonpatient samples in the general population of Rorschach dissertations. In this sense, the sample of Rorschach dissertations was "quasi-random" and unlikely to be biased in a way that would affect the findings of the present study. An alternative sampling strategy would have been to identify and purchase a random sample of dissertations (at $30 apiece) from Dissertation Abstracts International. However, this sampling strategy was deemed financially impractical and probably unnecessary. Later in this article, we report analyses suggesting that our sample of dissertations was unbiased.

Of the 83 dissertations that were manually searched, 66 (80%) did not fit the criteria for inclusion in the present review. Specifically, 34 (41% of 83) were eliminated because the subjects were children or adolescents, rather than adults; an additional 12 (14%) were eliminated because the subjects were prisoners, litigants, or members of another forensic group; an additional 13 (16%) were eliminated because the subjects were psychiatric or psychotherapy patients; and an additional 7 (8%) were eliminated either because the CS was not used or because the CS variables studied were not among the 14 variables included in the present review.

A total of 17 dissertations met criteria for inclusion in the present review. Of these, 7 (41%) had been authored either by a prominent Rorschach researcher (Hilsenroth, 1996/1997; Meyer, 1989/1991), or by a doctoral student working under the direct supervision of such a researcher (Burns, 1993/1994; DeLucas, 1997; Jacques, 1990/1991; Jansak, 1996/1997; Schiff, 1992/1993).

Thus, a total of 32 studies were included in the present study. These included 13 journal articles, 2 book chapters, and 17 dissertations. Table 1 provides information regarding these studies.

Selection of Variables

To avoid alpha inflation due to multiple statistical tests, we selected only a limited number of CS variables for inclusion in the present study. Fourteen Rorschach measures of distorted thinking or emotional problems were identified that (a) exhibited substantial discrepancies from the CS norms in the study by Shaffer et al. (1999) and (b) had

Table 1. Description of studies included in the present review

Study	Type of Publication	*N*	Participants	Comments
Alexander (1997/1998)	Dissertation	52	Black adults, ages 18–34	
Burns (1993/1994)	Dissertation	70	Normal married women	Same participants as Burns & Viglione (1996, 1997); dissertation advisor: D. Viglione
Caine et al. (1995)	Book chapter	20	Female undergraduates	Controls in malingering study
Calkins (1980/1981)	Dissertation	72	Undergraduates	
DeLucas (1997)	Dissertation	30	U.S. Navy and Air Force Security staff	Dissertation advisor: D. Viglione
Erstad (1995/1996)	Dissertation	12	Normal adults, ages 18–59	
		33	Normal adults, ages 61–95	
Frueh & Kinder (1994)	Article	20	Normal white male undergraduates	Controls in malingering study
Goldfinger (1998/1999)	Dissertation	21	Veterans and nonveterans without any psychiatric diagnosis	
Greenwald (1990)	Article	62	University students	
Hallet (1996)	Dissertation	126	Police detectives and patrol officers	
Hayslip et al. (1992)	Article	32	Community-residing older adults with normal color vision	
Hilsenroth (1996/1997)	Dissertation	50	Undergraduates	Same participants as Hilsenroth et al. (1997); present review used only variables not reported in that article.
Hilsenroth et al. (1997)	Article	50	Undergraduates	Same participants as Hilsenroth (1996/1997)
Jacques (1990/1991)	Dissertation	106	Community volunteers	Dissertation advisor: D. Viglione
Jansak (1996/1997)	Dissertation	30	Community volunteers without major psychiatric disorders	Dissertation advisor: D. Viglione
Kadle (1989)	Dissertation	30	Nonpatient adults, ages 63–87	
Kranau (1983/1984)	Dissertation	60	30 Hispanic Americans and 30 Anglo Americans	
Lipkin (1988/1989)	Dissertation	15	Nonpatient women, ages 55 or older	
Meisner (1988)	Article	29	UC-Berkeley graduate students	Controls in malingering study
Meyer (1989/1991)	Dissertation	265	Undergraduates	Same participants as Meyer (1992)
Netter & Viglione (1994)	Article	20	Nonpatients	Controls in malingering study
Paul (1987/1989)	Dissertation	60	Nonpatient adults, ages 65–94	
Perry & Kinder (1992)	Book chapter	20	White male undergraduates	Controls in malingering study
Ritzler & Nalesnik (1990)	Article	30	Hospital staff members, with no history of psychiatric hospitalization and no current symptoms	
Schiff (1992/1993)	Dissertation	25	M.D. and Ph.D. psychoanalytically oriented therapists	Dissertation advisor: P. Erdberg
Sloan et al. (1996)	Article	25	Marine reservists	
Smith et al. (1991)	Article	15	Undergraduate women	
Van Horn (1996)	Dissertation	30	Vietnam veterans	
Waehler (1991)	Article	28	Men ages 40–50 who have never been married	
Wald et al. (1990)	Article	28	Nonoffending mothers of incest victims	
Zacker (1997)	Article	53	Police applicants	
Zlotogorski et al. (1987)	Article	32	16 Hypnotically susceptible and 16 unsusceptible undergraduates	

been examined repeatedly in the 32 studies in the present review. The Rorschach variables that met these criteria were EB style (percent of ambitent protocols), Reflection responses, *X+%, X–%, Afr,* Form-Color responses (*FC*), Populars, *Sum Y,* the sum of Texture Responses (*Sum T*), *WSumC, MOR, WSum6,* Lambda, and the total number of Pure Human responses (*Pure H*). Many of these variables (e.g., EB, *X+%, X–%, WSum6, Populars, Sum Y, Sum T, WSumC,* Lambda) occupy a central role in Rorschach clinical interpretation and are among the most most commonly used CS scores. We predicted that these variables would show patterns in the independent adult

Table 2. Means and standard deviations of 13 Rorschach variables: A comparison of the present findings with Exner (1993) and Shaffer et al. (1999)

			Studies in Review						Studies in Review			
Variable	Exner Mean	Shaffer Mean	Mean	*d*	*k*	*N*	Exner SD	Shaffer SD	SD	Ratio	*k*	*N*
Reflections	0.08	0.46	0.57[a]	0.58	13	781	0.35	0.93	1.12[a]	3.2	11	701
X+%	0.79	0.51	0.60[ab]	−1.67	19	700	0.08	0.15	0.14[a]	1.7	17	608
X−%	0.07	0.21	0.19[a]	1.44	15	608	0.05	0.11	0.11[a]	2.2	15	608
Afr	0.69	0.48	0.49[a]	−1.14	17	745	0.16	0.20	0.19[a]	1.2	15	684
FC	4.09	1.76	1.82[a]	−1.29	17	916	1.88	1.91	1.67[a]	0.9	15	824
Populars	6.89	5.45	5.50[a]	−0.84	13	712	1.39	2.09	1.87[a]	1.3	10	590
Sum Y	0.57	1.37	1.74[a]	0.72	13	649	1.00	1.82	2.09[a]	2.1	10	558
Sum T	1.03	0.53	0.87[ab]	−0.18	16	799	0.58	0.83	1.07[ab]	1.8	13	708
WSumC	4.52	2.62	3.06[a]	−0.73	14	779	1.79	1.98	2.18[a]	1.1	9	540
MOR	0.70	1.05	1.07[a]	0.32	19	999	0.82	1.15	1.34[a]	1.6	14	806
WSum6	3.28	6.63	5.40[a]	0.41	11	382	2.89	7.99	7.78[a]	2.7	10	352
Lambda	0.58	1.22	0.88[a]	0.30	16	767	0.26	1.72	1.38[ab]	5.3	15	735
Pure H	3.40	2.67	2.12[a]	−0.74	8	303	1.80	2.16	1.59[b]	0.9	8	303

Notes: d = (mean of Exner norms − mean of studies in present review)/pooled standard deviation. Ratio = standard deviation in Exner norms: standard deviation in studies in present review. *k* = number of studies reporting this statistic. *N* = Number of participants in studies reporting this statistic.

[a]Statistically significant difference ($p < .001$) from statistic reported by Exner (1993).

[b]Statistically significant difference ($p < .001$) from statistic reported by Shaffer et al. (1999).

samples that were similar to those reported by Shaffer et al. (1999). Specifically, we hypothesized that compared with the CS norms, the independent samples of nonpatient adults would exhibit (a) a higher proportion of ambitent protocols, (b) a greater number of Reflection responses, (c) lower *X+%,* (d) higher *X–%,* (e) lower *Afr,* (f) lower *FC,* (g) fewer Popular responses, (h) higher *Sum Y,* (i), lower *Sum T,* (j) lower *WSumC,* (k) higher *MOR,* (l) higher *WSum6,* (m) higher Lambda, and (n) lower *Pure H.*

A few variables or types of variables from the study by Shaffer et al. (1999) were not included in the present analyses. First, the *SCZI* was not included because its mean and standard deviation have not been reported in recent normative tables (Exner, 1991, 1993). Second, a few variables (e.g., *FC: CF+C,* $T = 0$, Lambda $> .99$, *Pure H* < 2, *Dd*) were not included because they were seldom reported in the 32 studies in our sample. Whenever possible, closely related variables were used instead (e.g., *FC* instead of *FC: CF+C, Sum T* instead of $T = 0$, mean of Lambda instead of Lambda $> .99$, mean of *Pure H* instead of *Pure H* < 2).

COMPARISON OF ADULT NONPATIENT SAMPLES WITH CS NORMS AND SHAFFER ET AL. SAMPLE

Analyses

For 13 of the 14 Rorschach variables, the following procedures were followed. (a) Means and standard deviations were recorded for each sample that reported the variable (see Appendix A). (b) Weighted means for each variable were calculated by summing the means from each sample, weighting for sample size (see Table 2). (c) Weighted standard deviations for each variable were calculated by squaring the standard deviation from each sample to obtain a mean square, calculating the weighted mean of the mean squares (i.e., weighting for sample size), and then taking the square root of this quantity (see Table 2). (d) Independent sample *t* tests were performed to determine whether the means for the aggregated samples in Table 2 differed significantly from the CS norms for adult nonpatients (Exner, 1993, pp. 260–264) or from the figures reported by Shaffer et al. (1999). Because the standard deviations for the aggregated samples often apparently differed from the standard deviations listed in the normative tables, a conservative form of the *t* test that does not assume equal variances was used (Walpole & Myers, 1985). To limit the likelihood of Type I error due to multiple statistical tests, a highly conservative significance level of $p < .001$ was used for *t* tests. (e) *F* tests were performed to determine whether the standard deviations for the aggregated samples in Table 2 differed significantly from the standard deviations reported in the CS norms for adult nonpatients and in the study of Shaffer et al. (1999). Here, too, a significance level of $p < .001$ was used. These analyses were deemed important because clinicians routinely rely on the standard deviations reported in the CS norms to make judgments regarding the pathology of Rorschach protocols. (f) *T* tests and *F* tests were also performed for the

individual samples (see Appendix A), with comparisons with the CS norms only. If no standard deviation was available for a particular sample in Appendix A, then the standard deviation of the CS norms was used in performing the *t* tests.

Four of the Rorschach variables analyzed with *t* tests were highly positively skewed (i.e., Reflection responses, *Sum Y, Sum T, MOR*). However, because the aggregated samples for these four variables were all large ($N = 617–999$), the sampling distributions of their means would be practically normal by the central limit theorem, and therefore the results of the *t* tests were robust (Walpole & Myers, 1985).

For two variables (the proportion of ambitent protocols and the proportion of protocols with at least one Reflection response) *t* tests could not be performed because results were reported as proportions instead of means. For these two variables, the following procedures were followed. (a) The relevant proportions were recorded for each sample that reported the variable (see Appendixes B and C). (b) The proportions were pooled for each variable, weighting for sample size (see Table 3). (c) *Z* tests were performed (Walpole & Myers, 1985) to determine whether the proportions for the aggregated samples in Table 3 differed significantly from the proportions reported in the CS norms and in the study by Shaffer et al. (1999). A significance level of $p < .001$ was used for the *z* tests. (d) *Z* tests were also performed for the individual samples (see Appendixes B and C), with comparisons with CS norms only.

Results

As shown in Table 2, the means from the aggregated samples of nonpatient adults differed significantly from the CS norms (Exner, 1993) for all 13 variables, whereas the means from the aggregated samples significantly differed from the Shaffer et al. (1999) means for only two variables (*X+%* and *Sum T*). The standard deviations from the aggregated samples significantly differed from the CS norms for 12 of the 13 variables, whereas the standard deviations from the aggregated samples significantly differed from the Shaffer et al. standard deviations for three variables (*Sum T, Lambda,* and *Pure H*). Thus, as Table 2 indicates, the aggregated data from the 32 studies tended to differ significantly from the CS norms for nearly all analyses, but from the Shaffer et al. data for only a few analyses.

As shown in Table 3, a similar pattern emerged for the percentage of protocols that were ambitent or contained one or more Reflection responses. For both of these variables, the proportion of protocols in the aggregated samples differed significantly from the CS norms, but not from the findings of Shaffer et al. (1999). It should be noted that one variable (Reflections) was analyzed as both a continuous and a dichotomous variable, and thus appears in both Tables 2 and 3. Although most studies have treated Reflection responses as a continuous variable (i.e., reporting a mean and standard deviation), several studies have treated Reflections as dichotomous (i.e., zero Reflection responses versus one or more Reflection responses). Because studies have been inconsistent in their treatment of this variable, we included it in both Tables 2 and 3.

Thus, as predicted based on the findings of Shaffer et al. (1999), the aggregated samples of nonpatient adults exhibited (a) a higher proportion of ambitent protocols than the CS norms, (b) a greater number of Reflection responses, (c) lower *X+%,* (d) higher *X–%,* (e) lower *Afr,* (f) lower *FC,* (g) fewer Popular responses, (h) higher *Sum*

Table 3. Percentage of protocols that are ambitent or contain one or more reflection responses: A comparison of the present findings with Exner (1993) and Shaffer et al. (1999)

			Studies in Review				
Variable	**Exner**	**Shaffer**	**%**	***k***	***N***	**Dissertations**	**Published studies**
Ambitents	20	42	48[a]	9	290	51	44
Reflections	7	29	29[a]	8	368	29	—

Note: k = number of studies reporting this statistic. *N* = Number of participants in studies reporting this statistic.

[a]Statistically significant difference ($p < .001$) from statistic reported by Exner (1993).

[b]Statistically significant difference ($p < .001$) from statistic reported by Shaffer et al. (1999).

[c]Statistically significant difference ($p < .001$) between dissertations and published studies.

Y, (i) lower *Sum T,* (j) lower *WSumC,* (k) higher *MOR,* (l) higher *WSum6,* (m) higher Lambda,and (n) lower *Pure H.* All these findings were statistically significant. Thus, if Rorschach scores for a normal adult are interpreted using the CS norms, the adult will appear relatively self-focused and narcissistic (elevated Reflection scores), unconventional with impaired judgment and distorted perceptions of reality (low *X+%,* low Populars, high *X–%*), depressed, anxious, tense, and constrained in emotional expression (elevated Morbid responses, elevated *Sum Y,* low *WSumC*), insecure and fearful of involvement (elevated Lambda), vacillating and inefficient (elevated number of ambitents), with low empathy (low *Pure H*), a tendency to withdraw from emotions (low *Afr*), and poor emotional control (low *FC*). Results for each Rorschach variable in each sample are provided in Appendixes A, B, and C.

EXPLAINING DISCREPANCIES BETWEEN NONPATIENT ADULTS AND THE CS NORMS

For all 14 Rorschach variables in the present review, the aggregated means or proportions reported for nonpatient adults differed significantly in the predicted direction from the CS norms. Lipsey (1990) provided empirically based guidelines for interpreting the effect-size statistic *d,* which represents the standardized difference between means. According to these guidelines, the discrepancies of the nonpatient adults from the CS norms ranged from small ($d = -.18$) to very large ($d = -1.67$), and the median absolute discrepancy ($d = .73$) was large (see Table 2). In addition, significant differences in standard deviations were found for 12 out of 13 variables. The ratio of the standard deviation in the nonpatient samples to the standard deviation in the CS norms ranged from.9 to 5.3, and the median ratio was 1.7 (see Table 2). Although there are no generally accepted guidelines for interpreting these ratios, a median ratio of 1.7 indicates substantial and practically significant discrepancies between standard deviations. In combination, the discrepant standard deviations and discrepant means exert a multiplicative effect in distorting the interpretation of Rorschach scores. That is, if the standard deviation reported in the CS norms for a variable such as *X–%* is too small, and the mean is also too low, then the effect will be multiplicative, so that a very large number of nonpatient subjects will appear deviant on this variable.

How can such sizable discrepancies between the nonpatient adults and the CS norms be explained? Five possible explanations can be identified. We evaluate each in turn.[1]

Explanation 1: Administration or Scoring of Rorschachs Was Inadequate

The findings of the present review might be explained by suggesting that the administration or scoring of Rorschachs was poor in many or all of the 32 studies of nonpatient Americans. However, there is little or no evidence to support this speculation. To the contrary, administration and scoring appear to have been appropriate. Most of the 32 studies used trained administrators and scorers at the doctoral or advanced graduate levels and included reliability checks. Additionally, the authors of the 32 studies included prominent Rorschach researchers, including Bill Kinder, editor of the *Journal of Personality Assessment* (Caine, Frueh, & Kinder, 1995; Frueh & Kinder,1994; Perry & Kinder,1992), Robert Archer, editor of *Assessment* (Wald, Archer, & Winstead, 1990), Donald Viglione (Burns & Viglione, 1996; Netter and Viglione, 1994), Gregory Meyer (Meyer, 1989/1991), Mark Hilsenroth (Hilsenroth,1996/1997; Hilsenroth, Fowler, Padawer, & Handler, 1997; Sloan, Arsenault, Hilsenroth, Handler, & Harvill, 1996), Barry Ritzler (Ritzler & Nalesnik, 1990), and Charles Waehler (Waehler, 1991). Furthermore, several of the dissertations (Burns, 1993/1994; DeLucas, 1997; Jacques, 1990/1991; Jansak, 1996/1997; Schiff, 1992/1993) were supervised by Donald Viglione or Philip Erdberg, prominent researchers who have also served as instructors for the Rorschach Workshops. It is unlikely that so many well-known researchers failed to administer and score Rorschachs correctly. If they did, their failure would raise deeply disturbing problems of its own.[2]

Explanation 2: The Nonpatient Samples in the Present Review Are Unrepresentative of Normal American Adults

None of the various nonpatient samples included in the present review is representative of the entire American adult population. Specifically, none of the 32 studies attempted to obtain a true probability sample. Furthermore, nearly all the samples were distinctive in some way. Some groups were composed entirely of elderly adults, college students, blacks, Hispanics, psychoanalysts, veterans, or law enforcement personnel. However, although no single sample was representative of the American public, this fact does not explain the discrepancies from the CS

normative data. Indeed, the very diversity of the various groups actually tends to strengthen our conclusions. For example, as shown in Appendix A, an overabundance of Reflection responses, lowered *X+%* scores, and elevated *X–%* scores were observed in virtually all the samples reviewed, whether these samples were composed of college students, young black adults, veterans, or psychoanalysts. It is unlikely that the typical member of each of these groups is narcissistic or has poor reality testing. Because the findings were so consistent in such diverse populations, it is difficult to maintain that different results would have been discovered in a true probability sample of all American adults.

Explanation 3: The Dissertations Included in the Present Study Were a Biased Sample

Because the dissertations included in the present review were not randomly selected from among all Rorschach dissertations published over the past 25 years, the question arises whether they may have constituted a biased sample. Specifically, were the discrepancies from the CS norms, as reported in Tables 2 and 3, caused by an unrepresentative or biased sample of Rorschach dissertations? To examine this question, we compared the means and standard deviations for (a) the dissertations included in the present review, with (b) the published studies (i.e., articles and book chapters) in the present review, and (c) the CS norms (Exner, 1993).

Table 4 shows the results for continuous variables, and Table 3 shows the results for variables reported as proportions. To control for alpha inflation due to multiple tests, the level of statistical significance was set at $p = .0036$ using the Bonferroni correction (.05/14). As may be seen, for 10 out of 14 CS variables, there was no significant difference between the mean scores or proportions for dissertations versus published studies. For 3 of the 4 significant differences, the published studies were actually more discrepant from the CS norms than were the dissertations. Specifically, the published studies were significantly lower on *X+%,* higher on *X–%,* and lower on *Sum T* than were the dissertations. Finally, for the remaining significant difference, the dissertations exhibited a significantly higher mean level of Reflection responses (.63) than did the published studies (.32). However, even in this case, the published studies' mean of .32 was still four times as high as the mean of .08 reported in the CS norms (Exner, 1993).

Table 4. Means and standard deviations for the Exner (1993) norms versus dissertations (D) and published studies (PS) in the present review

	Means			Standard Deviations		
Variable	**Exner**	**D**	**PS**	**Exner**	**D**	**PS**
Reflections	0.08	0.63[a]	0.32[a]	0.35	1.13	1.05
X+%	0.79	0.64[a]	0.55[a]	0.08	0.14	0.14
X−%	0.07	0.18[a]	0.21[a]	0.05	0.12	0.11
Afr	0.69	0.49	0.48	0.16	0.19[a]	0.16[a]
FC	4.09	1.76	2.03	1.88	1.69	1.59
Populars	6.89	5.57	5.31	1.39	1.89	1.76
Sum Y	0.57	1.81	1.55	1.00	2.15[a]	1.69[a]
Sum T	1.03	1.00[a]	0.64[a]	0.58	1.14[a]	0.85[a]
WSumC	4.52	3.08	2.94	1.79	2.21	2.01
MOR	0.70	1.10	0.95	0.82	1.34	1.30
WSum6	3.28	5.96	4.54	2.89	7.99	7.38
Lambda	0.58	0.78	1.08	0.26	0.90[a]	2.11[a]
Pure H	3.40	2.18	1.93	1.80	1.55	1.72

[a]Statistically significant difference ($p < .0036$) between dissertations and published studies.

Thus, the numbers in Table 4 do not indicate that the dissertations in the present review were a biased or unusual sample. Overall, the means and proportions reported for dissertations resemble the means and proportions for published studies. Although a few significant differences appear, they usually occur because the published studies are even more discrepant from the CS norms than are the dissertations. The results of the present review cannot be explained as due to a biased or unusual sample of dissertations.

Explanation 4: The CS Normative Data Were Accurate in the 1970s, but the Rorschach Scores of Americans Have Changed in the Intervening Years

It might be argued that although the CS norms were accurate when the normative data were first collected in the 1970s, Americans have changed during the intervening years, thereby creating the discrepancies observed in the present study. However, such an explanation does not accord with the available evidence.

First, the suggestion that Americans' Rorschach scores have drifted away from the CS norms over the past few decades is based on the assumption that the current normative data for the CS (Exner, 1993) were gathered in the 1970s. This assumption appears to be mistaken. Although recent editions of *TRACS* (Exner, 1991, 1993) do not state the specific dates that the CS normative data were gathered, Exner reported in 1993 that "these data have accumulated over a period of more than 20 years"

Table 5. Comprehensive System norms from the 1970s to the 1990s (Exner, 1974, 1978, 1986, 1993): A comparison of means with Shaffer et al. (1999) and the present findings

	Exner					
Variable	**1974**	**1978**	**1986**	**1993**	**Shaffer**	**Present Studies**
Reflections	0.20	0.14	0.12	0.08	0.46	0.57
X+%	0.84	0.81	0.80	0.79	0.51	0.60
X−%	—	—	0.06	0.07	0.21	0.19
Afr	—	0.69	0.66	0.69	0.48	0.49
FC	3.20	3.56	3.87	4.09	1.76	1.82
Populars	6.70	6.45	6.66	6.89	5.45	5.50
Sum Y	1.30	1.11	0.98	0.57	1.37	1.74
Sum T	1.40	1.18	1.16	1.03	0.53	0.87
WSumC	3.60	3.73	4.23	4.52	2.62	3.06
MOR	—	—	0.70	0.70	1.05	1.07
WSum6	—	—	3.96	3.28	6.63	5.40
Lambda	0.74	0.82	0.59	0.58	1.22	0.88
Pure H	—	—	3.07	3.40	2.67	2.12

(p. 258). Thus, the CS norms apparently do not simply represent data from the 1970s, but instead are based on data collected over a period of two decades, from approximately 1973 to 1993. In fact, it appears that many protocols in the normative pool may have been gathered at approximately the same time when several studies in the present review were published.

Second, as can be seen in Table 5, the adult normative data for the CS published in *TRACS* from the 1970s to the 1990s (Exner, 1974, 1978, 1986, 1993) provide limited evidence that some CS scores may have shifted over the past 25 years. However, the patterns of change in Table 5 generally cannot explain the discrepancies shown in Tables 2 and 3. For example, according to the normative data, the mean number of Reflection responses decreased from .20 in 1974 to .08 in 1993. But if Reflection responses have become less frequent over time, this trend cannot explain why the studies included in the present article report a very high level of Reflection responses (mean =.59). Similarly, according to the normative data, the mean of *Sum Y* decreased from 1.30 in 1974 to .57 in 1993. But if *Sum Y* responses have become less frequent over time, this finding cannot explain why the studies in the present review found a very high level of *Sum Y* (mean = 1.74). Similarly, norms for *FC, WSumC,* and Lambda have shifted over time in the opposite direction from the discrepancies shown in Tables 2 and 3.

Third, even aside from the CS normative data, findings from several early studies indicate that some Rorschach scores have remained stable over the years. For example, in the first published group study of Reflection responses, Exner (1969) reported that 35% of protocols among normal college students contained at least one Reflection response (see Table 6).[3] This 35% figure from the 1960s is actually slightly higher than the average 29% figure reported in normal adult samples in the 1980s and 1990s (see Table 3). Similar findings are reported in studies from the 1970s. For example, Winter and Exner (1973; cited by Exner, 1993) found that 39% of protocols from performing artists contained a Reflection response, and Exner, Weiss, Coleman, and Rose (1979; cited by Exner, 1993) found a rate of 36% among theatrical dancers (see Table 6). Although *TRACS* (Exner, 1993) interprets these numbers as evidence of narcissism, in fact, the proportion of Reflection responses in these studies was similar to what has been found in samples of normal American adults over the decades (see Table 3). For example, a relatively early study of undergraduates (Calkins, 1980/1981) found at least one Reflection response in 39% of protocols.

There is a fourth reason to doubt that Americans' CS scores have changed dramatically over time. Although the 32 studies in the present article extend from the late 1970s to the late 1990s, their findings do not indicate that most CS scores have changed during this time period. Table 7 shows the correlations between the year each study was published (or each dissertation was completed) and the mean value that the study reported for CS scores. As can be seen, only 4 of the 14 Rorschach variables (*X–%, Afr, FC,* Lambda) showed significant changes over time, even

Table 6. Percentage of protocols with at least one Reflection response: Exner (1993) norms as compared with other adult samples reported by Exner

	N	**% Protocols**
Norms (Exner, 1993)	700	7
Exner (1969)		
College students	20	35
Winter & Exner (1973); cited by Exner, 1993)		
Successful performing artists with no psychiatric history	18	39
Exner et al. (1979; cited by Exner, 1993)		
Successful theatrical dancers	39	36
Exner (1993)		
Male candidates for engineering job	20	30
Exner (1993)		
Clergy	?	30
Surgeons	?	24
Theatrical employees	?	20

Table 7. Correlations between the year a study was published (or dissertation was completed) and the mean of the Rorschach variable in that study

Variable	No. of Studies	Correlation (Spearman's ρ)	*p*
Reflections (mean)	13	.01	.98
Reflections (% protocols)	8	−.07	.86
X+% (mean)	19	−.17	.49
X−% (mean)	15	−.56	.03
Afr (mean)	17	−.49	.04
FC (mean)	17	−.51	.04
Populars (mean)	13	.08	.79
Sum Y (mean)	13	.03	.91
Sum T (mean)	16	−.20	.46
WSumC (mean)	14	−.22	.45
MOR (mean)	19	.28	.24
WSum6 (mean)	11	.49	.12
Lambda (mean)	16	.55	.03
Pure H (mean)	8	−.26	.52
EB (% ambitent)	9	−.60	.09

when the level of statistical significance was set liberally at .05.[4]

Even for these four variables, the observed time trends do not neatly explain the discrepancies observed in the present study. For example, the correlation between *X–%* and year of study was negative ($r = -.51$). In other words, the trend over time has been for *X–%* scores to decrease, and thus come closer to the normative data (Exner, 1993). This pattern of change contradicts the hypothesis that Rorschach scores have drifted from the 1970s levels. As another example, mean *Afr* scores from studies in the 1980s tended to cluster between .50 and .60 (substantially below normative data from the 1970s, 1980s, and 1990s), whereas *Afr* scores from studies in the 1990s tended to cluster even lower, between .40 and .50. Thus, the pattern was not for *Afr* scores to be close to norms at the start and then to drift. Rather, *Afr* scores in the 1980s were discrepant from the 1980s norms, and *Afr* scores in the 1990s became even more discrepant from the 1990s norms (see also Loucks et al., 1980).

How are the changes over time in *X–%, Afr, FC,* and Lambda to be explained? At least in part, the shifts in scores may be due to changes in the scoring rules for the CS over the past 25 years. The CS scoring rules have repeatedly been changed and refined in scoring manuals published by the Rorschach Workshops (Exner, 1985, 1990, 1995b; Exner, Weiner, & Schyler, 1976). It seems plausible that these changes have resulted in changes in means and standard deviations for at least some CS scores. In addition, during the 1980s, protocols with 13 or fewer responses were dropped from the Rorschach Workshops subject pool, and new norms were calculated (Exner, 1991). It is possible that this alteration in the subject pool contributed to the changes in *X–%, Afr, FC* and Lambda.

Explanation 5: The Norms for Important CS Scores Do Not Currently Represent American Nonpatient Adults and Probably Never Did

As has been shown, the discrepancies between the CS norms and the studies in the present review cannot be dismissed as due to shifts in Americans' Rorschach scores over time or to methodological artifacts of the present review. Thus, a simple conclusion suggests itself: The CS norms for the variables in the present study do not currently represent American nonpatient adults and probably never did. In light of data from several sources, this straightforward explanation seems much more plausible than the alternatives.

If there are errors in the CS norms, how did they come about? Two explanations seem plausible. First, it is possible that the administration or scoring of Rorschach protocols for the CS normative pool was problematic. However, information provided in *TRACS* regarding administration and scoring is too vague to allow a thorough evaluation. For example, TRACS provides little detail regarding (a) the qualifications of the administrators and scorers or how they were recruited, (b) the specific sites where data were collected, (c) the number of protocols collected at each site, (d) the exact years when protocols were collected, or (e) procedures that were instituted to ensure integrity of test administration and recording. As we have reported elsewhere (Garb et al., in press; Wood & Lilienfeld, 1999; Wood et al., 1996a, 1996b), the unpublished Rorschach Workshops studies that form the empirical foundation for the CS are often unavailable for scrutiny by independent scholars. It is unfortunate that these studies cannot be examined because they might provide more detail regarding possible scoring or administration problems.

A second explanation for the shortcomings of the CS norms can be suggested. Perhaps the original sample was not truly representative of nonpatient Americans, but instead was a "super normal" group that excluded all subjects with any form of psychopathology. There is some limited evidence to support such a conclusion: *TRACS* states that no members of the normative sample had any admitted psychiatric history (Exner, 1993, p. 258). Fur-

thermore, 172 of the 700 adult subject volunteered through social or interest organizations such as the PTA or Audubon groups. Arguably, such factors might have caused the normative sample to be super normal and unrepresentative of all American nonpatients.

However, such an explanation can probably be rejected for several reasons. First, *TRACS* nowhere claims that the CS normative sample is super normal or exceptionally healthy compared with most nonpatient Americans. On the contrary, Exner (1991) has explicitly warned that some nonpatients in the CS normative sample are not even minimally normal:

> *The only element common to all subjects is the absence of psychiatric history. The subjects are not necessarily normal; they are simply not patients. As such, they represent a vast array of individual differences, with dimensions ranging from introversive to extratensive, well controlled to poorly controlled, gregarious to isolated, strange to sturdy, and so on. (pp. 460–461)*

Exner's (1991) forthright statement that the subjects in the normative sample were "not necessarily normal," and that they ranged from "well controlled to poorly controlled . . . strange to sturdy," contradicts any claim that the sample was super normal.

In addition, the procedures used to select the CS normative sample apparently were never intended to identify a super normal group. For example, the psychiatric background of subjects seems to have been determined entirely by retrospective self-report (Exner, 1993, p. 258), although many disturbed individuals either lack a clear-cut or documented psychiatric history or fail to report it. Furthermore, there is no indication in *TRACS* that subjects were given detailed psychodiagnostic interviews to rule out the presence of Axis I or Axis II psychopathology or that they were asked about their legal or criminal histories.

Any claim that the CS normative sample was super normal has two additional weaknesses. First, even if the normative sample were somewhat above normal, this fact could not account for the sizable mean discrepancies reported in the present review. As may be seen in Table 2, the between-group discrepancies were large for Reflections ($d = 0.58$), Populars ($d = -0.84$), *Sum Y* ($d = 0.72$), *WSumC* ($d = -0.73$), and *Pure H* ($d = -0.74$), and very large for *X+%* ($d = -1.67$), *X–%* ($d = 1.44$), *Afr* ($d = -1.14$), and *FC* ($d = -1.29$). These discrepancies are too large to be accounted for entirely by a somewhat above-average normative group. For example, the effect size of –1.67 for *X+%* indicates that the mean of the nonpatients in the present review was below the 5th percentile for the CS normative sample.

In addition, the discrepancies in standard deviations cannot be explained by arguing that the normative group was above average. The nonpatient adults in the present review show substantially more variability than the CS norms. For five variables, the ratio of standard deviations is more than 2:1, including Lambda (5.3:1), Reflections (3.2:1), *WSum6* (2.7:1), *X–%* (2.2:1), and *Sum Y* (2.1:1). Discrepancies this large cannot be readily explained by arguing that the CS normative sample was unusually homogeneous. Indeed, *TRACS* presents extensive data regarding the educational, racial, socioeconomic, and geographical composition of the normative sample to demonstrate that it was generally not more homogeneous than the U.S. population (Exner, 1993, pp. 258–262).

VALIDITY OF THE COMPREHENSIVE SYSTEM: A REASSESSMENT OF COMMON KNOWLEDGE IN LIGHT OF THE PRESENT FINDINGS

If the CS norms are in error, then much "common knowledge" about CS validity is also likely to be in error. Important claims about the validity of CS scores have often been established by comparing patient or participant groups with the CS norms. Such claims must be reevaluated critically in light of the apparent problems with the norms.

For example, Gacono and Meloy (1992, 1994) compared prisoners with antisocial personality disorder (ASPD) with the CS norms and concluded that the prisoners' Rorschachs 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 Tables 2 and 3 show, the same pattern of Rorschach scores is shown generally by nonpatient American adults. That is, compared with the CS norms, normal Americans also have a high proportion of Reflections, low *Sum T,* low *FC,* high *Sum Y,* low *X+%,* high *X–%* and high *WSum6.* Gacono and Meloy's findings have been held forth as evidence that the Rorschach validly measures aspects of

Table 8. Mean validity coefficients for studies in the meta-analysis by Hiller et al. (1999) by type of Rorschach system (CS or Non-CS) and whether norms were used to calculate validity coefficients

			Mean Validity Coefficient (*r*)	
Type of system	**No. of Studies**	**No. of Subjects**	**Weighted**	**Unweighted**
CS with norms	4	134	.34	.56
CS without norms	7	375	.24	.27
Non-CS	17	994	.25	.22

Notes: CS with norms: Comprehensive System studies in which validity coefficients were calculated with normative data. CS without norms: Comprehensive System studies in which validity coefficients were calculated without normative data. Non-CS: Studies that did not not use the Comprehensive System.

ASPD or psychopathy (Kubiszyn et al., 2000; Weiner, 1997). However, a much different interpretation is possible: The CS norms may often make adults look pathological, whether they are normal Americans or psychopathic prisoners with ASPD.

A second example is provided in a recent literature review by Viglione (1999), who argues that diffuse shading responses (*Sum Y*) are related to stress and anxiety. In support of this contention, Viglione's article cites four studies of traumatized patients that have reported elevated *Sum Y* in comparison with the CS norms (Hartman et al., 1990; Kaser-Boyd, 1993; Sloan, Arsenault, Hilsenroth, Harvill, & Handler, 1995; Swanson, Blount, & Bruno, 1990). However, as can be seen in Table 2, normal Americans also tend to show elevated *Sum Y.* In order to determine whether *Sum Y* is related to traumatic stress, researchers need to collect data from appropriate comparison groups (e.g., nontraumatized patients). In fact, studies using comparison groups have not generally found an elevation of *Sum Y* in posttraumatic stress disorder patients compared with other patient or nonpatient groups (Frueh & Kinder, 1994; Goldfinger, 1998/1999; Van Horn, 1996; see review by Wood et al., 2000a).

A third example of the need for reevaluation is provided by a recent meta-analysis of Rorschach and Minnesota Multiphasic Personality Inventory (MMPI) studies by Hiller, Rosenthal, Bornstein, Berry, and Brunell-Neuleib (1999). The Hiller et al. meta-analysis included data from four Rorschach studies in which comparisons were made with the CS norms rather than with true comparison groups (Cruz, Brier, & Reznikoff, 1997; Exner, Colligan, Boll, Stischer, & Hillman, 1996; Kaser-Boyd, 1993; Zimmerman & Dillard, 1994). These four studies included several variables that appear in Tables 2 and 3 of the present article, such as *X+%, X–%, WSum6,* and Populars. The incorporation of these four studies in the meta-analysis was problematic, however, because the use of the CS normative data may have led to inflated estimates for the validity coefficients. For example, Table 8 shows the mean validity coefficients for Rorschach studies in the Hiller et al. meta-analysis, grouped by CS studies that used the norms to compute validity coefficients, CS studies that did not use the norms to compute validity coefficients, and non-CS studies. Validity coefficients were similar for non-CS studies and CS studies that did not use norms (weighted validity coefficients were $r = .25$ and .24, respectively, and unweighted validity coefficients were $r = .22$ and .27, respectively). However, the CS studies that used comparisons with normative data had higher weighted ($r = .34$) and unweighted ($r = .56$) mean validity coefficients. The possibility arises that by using the problematic CS norms to calculate validity coefficients, Hiller et al. may have inadvertently inflated the overall validity coefficients for the Rorschach in their meta-analysis.

Many similar examples can be identified. For instance, relying on implicit or explicit comparisons with the CS normative data, Rorschach researchers have published studies regarding the supposedly deviant scores of children with learning disabilities (Acklin, 1990; Champion, Doughtie, Johnson & McCreary, 1984; Cruz et al.,1997), sexual abuse victims (Zimmerman & Dillard, 1994), burn patients (Holaday, 1998; Holaday & Whittenberg, 1994), police applicants (Zacker, 1997), airline pilots treated for alcohol or substance abuse (Ganellen, 1994), mental health professionals who engaged in sexualized dual relationships (Celenza & Hilsenroth, 1997), and pedophiles (Bridges, Wilson, & Gacono, 1998). Because these studies relied on comparisons with the problematic CS norms, their conclusions must also be regarded as problematic. As

the present study shows, individuals who take the Rorschach may differ from the CS norms, not because the individuals have truly deviant scores on the test, but because the norms tend to make even normal individuals appear more pathological than they really are.

IMPLICATIONS FOR CLINICAL USE OF THE COMPREHENSIVE SYSTEM

If the CS norms do not accurately represent the Rorschach scores of nonpatient Americans, what should psychologists do who use the CS in clinical or forensic work? One tempting option is to reject the present findings as inconsistent with first-hand clinical experience. Both Finn (1996) and Weiner (1999) reported a first-hand clinical experience that many other psychologists have probably shared. The MMPI-2 and the Rorschach are administered to a troubled but defensive client. The MMPI-2 profile looks normal, whereas several Rorschach scores indicate the presence of psychopathology. Many psychologists who have had this experience are likely to conclude, as did Finn and Weiner, that the Rorschach is a sensitive clinical tool that is capable of detecting hidden psychopathology that the MMPI-2 misses. However, the findings of the present review suggest a much different interpretation. If a patient looks pathological on the Rorschach but not on the MMPI-2, the reason may not be that the Rorschach is a more sensitive instrument than the MMPI-2, but that the norms of the former are wrong. The CS norms tend to make clients look more pathological than they really are, whether they are healthy or unhealthy.

It is also tempting to dismiss the discrepancies reported here as too small to make a practical difference and continue using the CS norms as before. Nevertheless, the large median effect size we have reported between the means derived from CS norms and the means from other CS studies renders this argument untenable. Moreover, such a course would be misguided because the apparent problems with the CS norms bear important practical implications for Rorschach interpretation.

Data presented by Shaffer et al. (1999, Table 6, pp. 313–314) show how the kinds of discrepancies identified in the present review can dramatically affect Rorschach interpretations and make normal individuals appear pathological. For example, according to *TRACS, X–%* scores greater than .20 occur in less than 1% of the CS normative sample and indicate "significant problems that promote perceptual inaccuracy and/or mediational distortion" (Exner, 1991, p. 214). Yet in Shaffer's sample of nonpatient adults, 48% had *X–%* scores greater than .20. As another example, Rorschach protocols with zero Texture responses are said to be relatively rare, occurring in only 11% of the CS normative sample. Accordingly, Texture-less Rorschachs supposedly indicate that the individual is "guarded and/or distant in interpersonal contacts" (Exner, 1993, p. 385). Yet in Shaffer's sample of nonpatient adults, 64% of protocols were Texture-less. Thus, use of the CS norms for Texture could lead to a mistaken conclusion that nearly two-thirds of nonpatient adults are guarded and/or distant.

As a third example, according to the CS norms, an *Afr* below .44 is extremely deviant, and usually indicates that a person is "quite uncomfortable around emotion," and often suggests that the person is "socially constrained or even isolated" (Exner, 1991, p. 1998). Yet in the nonpatient sample of Shaffer et al., 33% of participants had *Afr* below .40. Thus, about one third of nonpatient adults may appear emotionally maladjusted if compared to the CS norms. As a fourth example, according to Exner, Reflection responses "are not expected to appear in the records of older adolescents or adults" (p. 149). Even a single Reflection response supposedly indicates that "a nuclear element in the subject's self-image is a narcissistic-like feature that includes a marked tendency to overvalue personal worth" (p. 173). Yet in the Shaffer et al. sample, as well as in the aggregated data in the present review, 29% of protocols contained at least one Reflection response. It is sobering to realize that in forensic contexts such as custody disputes, nearly one third of normal parents are likely to be characterized as narcissistic by the CS interpretive rules.

Psychologists who work in clinical and forensic settings cannot avoid such problems by the simple expedient of combining Rorschach results with other sources of information, such as clinical interviews, MMPI-2 results, and psychiatric history. Studies have shown that clinical judgments can become less accurate if invalid test results are combined with more valid sources of information (Garb, 1984, 1998; Wiggins, 1988). Moreover, the question arises why the CS and its problematic norms should be used, when valid alternative tests with well-established norms are readily available (e.g., the MMPI-2). Whether the Rorschach is interpreted by itself or in combination with other data, the possibility of overdiagnosing psychopathology is likely to increase if the CS norms are used.

These problems do not appear to be limited to the 14 Rorschach variables in the present review. For example, Shaffer et al. (1999) found that 16% of normal adults scored in the pathological range (> 4) on the Schizophrenia Index (compared with less than 1% of the CS normative sample), and 15% scored in the pathological range (> 5) on the Depression Index (compared with less than 4% of the normative sample). Nor does it appear that such problems are limited to the adult norms alone. Wideman (1998) found that 30 gifted children (cognitive ability scores > 125) and 30 nongifted children (cognitive ability scores 90–109) appeared maladjusted for 12 (86%) out of 14 Rorschach variables when compared with the CS norms. Other researchers have found that normal children often look pathological compared with the CS norms (DeSousa, 1993; Ebert, 1991/1992; Esmail, 1996/1997; Harper & Scott, 1990; Kelly & Ben-Meir, 1993; Leifer, Shapiro, Martone, & Kassem, 1991; Salyer, Holmstrom, & Noshpitz, 1991; Yanofsky, 1994/1995; see also Loucks et al., 1980).[5] Thus, although we examined only adult samples, our findings appear to extend to children as well. In fact, the conclusions of the present study were anticipated 7 years ago by Kelly and Ben-Meir (1993), who discovered similar discrepancies in the Rorschach scores of nonpatient children. These researchers suggested the possibility that the CS norms

do not provide us with the most accurate prediction of what most "normal" children will score when tested by most trained professionals. If that is true, the potential for overpathologizing children on the basis of their Rorschach responses gives cause for concern. Moreover, since most clinicians and researchers do not use control groups, they do not have a self-correction mechanism in place; that is, they will never know that they are overpathologizing. (p. 112)

In light of the present findings, we recommend that psychologists not use the present CS norms in clinical or forensic work, with either children or adults. Psychologists who use these norms run the risk of attaching false and negative labels to clients and thereby potentially harming them. Ethical principles require that psychologists take reasonable steps to avoid harming clients or patients (American Psychological Association [APA], 1992, Standard 1.14; see also APA, 1999). Professional standards also require that psychologists use tests or techniques in a manner that is appropriate in light of relevant research (Standard 2.02a) and avoid using obsolete tests and measures (Standard 2.07b).

Some psychologists may be reluctant to accept such a recommendation and instead may continue to use the CS norms in clinical or forensic settings. If so, ethical principles seem to require that when communicating their findings, psychologists should forthrightly describe the limitations of the CS norms (APA Standards 2.08a, 7.04b).

IMPLICATIONS FOR RESEARCH ON THE COMPREHENSIVE SYSTEM

The present findings have important implications for researchers who examine the CS. First and most obviously, Rorschach validity studies in the future must be careful to gather data from appropriate comparison groups rather than rely on comparisons to the CS norms. The common research practice of using CS normative data in lieu of true comparison groups has repeatedly been criticized by Rorschach scholars (Exner, Kinder, & Curtiss, 1995; Garb et al., in press; Ritzler & Exner, 1995; Viglione, 1997; Viglione & Exner, 1995; Weiner, 1995; Wood et al., 1999, Wood et al., 2000a). Exner's position on this issue appears to have been ambiguous. On the one hand, he has stated, "Occasionally, researchers use published norms as a control sample against which comparisons for small groups are made. This tactic is naive at best, and invariably leads to faulty and misleading conclusions" (Exner, 1991, p. 460). On the other hand, he has endorsed the practice of making comparisons to norms if the discrepancies are large and statistical tests are not performed:

In some instances, it may be useful to call attention to normative data, especially when a group deviates markedly for some features. For example, a review of tables of data (Exner, 1991) reveals that only 1% of the 700 subjects in the nonpatient sample have X–% (Minus total responses) greater than 20%, compared with the 90% of the 320 subjects in the reference sample of schizophrenics. The data are obvious, and no statistical manipulations would enhance the meaning. (Exner & Sendin, 1997)

Here we have introduced a new reason that researchers should avoid making comparisons with the CS norms: The norms themselves may be inaccurate. Researchers who make comparisons with the CS norms may discover effects where there are none. As discussed earlier, past studies of prisoners with ASPD, veterans with posttraumatic stress disorder, or abused children may have been mistaken in their conclusions because they relied on comparisons with the CS norms.

A second implication of the present findings is that Rorschach researchers need to reconduct past studies that have relied on comparisons with the CS normative data. For example, it would be a mistake for researchers to assume that Gacono and Meloy's (1992, 1994) studies have empirically validated Rorschach scores as indicators of ASPD. New studies with appropriate comparison groups and better methodology are necessary to ascertain the relationship of ASPD to Rorschach variables.

A third implication of the present findings is that in the future, meta-analysts should either exclude or separately analyze Rorschach studies that have used comparisons with normative data. Inclusion of such studies in a meta-analysis can have the effect of inflating mean validity coefficients. Past meta-analyses that included such studies (e.g., Hiller et al., 1999) need to be interpreted in light of this limitation.

Fourth, researchers may wish to establish a new set of national norms for the CS or develop local norms for specific groups of participants, such as Texas Hispanics or older adults. However, attempts to develop local norms for the CS face several obstacles and may be inadvisable. First, if the means and standard deviations for such norms are to be stable, then a substantial number of participants must be included in the normative sample. As can be seen in Appendix A, means and standard deviations can vary considerably if samples are small. A single, national normative study could be more economical than several separate studies to establish local norms.

Another problem with establishing local norms concerns the quality of test administration and scoring. About half of the CS variables have scoring reliability below .85 using the intraclass correlation coefficient (Acklin, McDowell, Verschell, & Chan, 2000). If local norms are collected for Texas Hispanics, for example, and those norms differ from national or other norms, the question will arise, does this difference reflect a true difference between groups, or is it simply due to differences in scoring or test administration? In many cases, such a question can be addressed only by replication studies that use the same population of participants but different test administrators and scorers.

Local norms can also lead to problems of interpretation. For example, Kranau (1983/1984) found that Hispanics had significantly higher *X+%* than Anglos. But consider the case of a Hispanic who scores at the mean level of *X+%* for Anglos. Is such a score to be considered "average" (according to the Anglo norms) or "below average" (according to the Hispanic norms)? Furthermore, if a Hispanic and an Anglo both have the same *X+%* score, do their scores have the same or different meanings for purposes of diagnosis or prediction? Such questions apply to the cross-cultural use of any test, but they become particularly salient when local norms are used.

Attempts to develop local norms for the CS still lie in the future, however. At present, the CS seems to lack either local norms or accurate national norms. Such a situation presents grave difficulties for clinicians and researchers who use the Comprehensive System for the Rorschach.

APPENDIX A: NORMS (EXNER, 1993) FOR MEANS OF 13 RORSCHACH VARIABLES AS COMPARED WITH INDEPENDENT SAMPLES OF NONPATIENT ADULTS

Variable/Study	*N*	Mean (SD)	*t*	*p*	*F*	*p*
Reflections						
Norms	700	.08 (.35)				
Calkins (1980/1981)	72	.76 (1.32)	4.36	.00004	14.22	<.00001
DeLucas (1997)	30	.48 (.85)	2.57	.01555	5.90	<.00001
Erstad (1995/1996)	12	.75 (.97)	2.39	.03576	7.68	<.00001
Erstad (1995/1996)	33	.18 (.46)	1.23	.22642	1.73	.00817
Hilsenroth et al. (1997)	50	.30 —	4.29	.00007	—	—
Jacques (1990)	106	.66 (1.13)	5.25	<.00001	10.42	<.00001
Kadle (1989)	30	.47 (1.61)	1.32	.19533	21.16	<.00001
Meyer (1989/1991)	265	.79 (1.19)	9.56	<.00001	11.56	<.00001
Paul (1987/1989)	60	.13 (.47)	.80	.42369	1.80	.00034
Ritzler & Nalesnik (1990)	30	.10 —	.31	.76125	—	—
Schiff (1992/1993)	25	.60 (1.26)	2.06	.05027	12.96	<.00001
Smith et al. (1991)	15	.27 (.46)	1.59	.13365	1.73	.04612
Zacker (1997)	53	.49 (1.17)	2.54	.01398	11.17	<.00001
Total (except norms)	781	.57 (1.12)	11.61	<.00001	10.24	<.00001
X+%						
Norms	700	.79 (.08)				

APPENDIX A: *Continued*

Variable/Study	*N*	Mean (SD)	*t*	*p*	*F*	*p*
DeLucas (1997)	30	.57 (.15)	−7.98	<.00001	3.52	<.00001
Erstad (1995/1996)	12	.53 (.13)	−6.90	.00002	2.64	.00258
Erstad (1995/1996)	33	.60 (.12)	−9.00	<.00001	2.25	.00012
Frueh & Kinder (1994)	20	.52 (.14)	−8.58	<.00001	3.06	.00001
Goldfinger (1998/1999)	21	.47 (.12)	−12.14	<.00001	2.25	.00143
Greenwald (1990)	62	.52 (.13)	−16.09	<.00001	2.64	<.00001
Hallett (1996)	126	.68 (.14)	−8.57	<.00001	3.06	<.00001
Kadle(1989)	30	.71 (.10)	−4.32	.00015	1.56	.03115
Kranau (1983/1984)	60	.78	−.93	.35598	—	—
Lipkin (1988/1989)	15	.58 (.14)	−5.79	.00004	3.06	.00013
Netter & Viglione (1994)	20	.55 (.15)	−7.13	<.00001	3.52	<.00001
Paul (1987/1989)	60	.54 (.16)	−11.97	<.00001	4.00	<.00001
Perry & Kinder (1992)	20	.50 (.10)	−12.85	<.00001	1.56	.05958
Schiff (1992/1993)	25	.55 (.12)	−9.92	<.00001	2.25	.00061
Sloan et al. (1996)	25	.57 (.18)	−6.09	<.00001	5.06	<.00001
Waehler (1991)	28	.53 (.16)	−8.56	<.00001	4.00	<.00001
Wald et al. (1990)	28	.55 (.16)	−7.90	<.00001	4.00	<.00001
Zacker (1997)	53	.60 (.12)	−11.34	<.00001	2.25	<.00001
Zlotogorski et al. (1987)	32	.59	−13.83	<.00001	—	—
Total (except norms)	700	.60 (.14)	−31.18	<.00001	3.06	<.00001
X−%						
Norms	700	.07 (.05)				
Burns (1993/1994)	70	.22 (.14)	8.91	<.00001	7.84	<.00001
DeLucas (1997)	30	.20 (.13)	5.46	.00001	6.76	<.00001
Erstad (1995/1996)	12	.11 (.10)	1.38	.19396	4.00	.00001
Erstad (1995/1996)	33	.14 (.09)	4.44	.00010	3.24	<.00001
Frueh & Kinder (1994)	20	.17 (.12)	3.72	.00144	5.76	<.00001
Goldfinger (1998/1999)	21	.19 (.12)	4.57	.00018	5.76	<.00001
Greenwald (1990)	62	.21 (.10)	10.90	<.00001	4.00	<.00001
Hallett (1996)	126	.14 (.10)	7.69	<.00001	4.00	<.00001
Netter & Viglione (1994)	20	.20 (.13)	4.46	.00026	6.76	<.00001
Paul (1987/1989)	60	.26 (.14)	10.46	<.00001	7.84	<.00001
Perry & Kinder (1992)	20	.17 (.08)	5.56	.00002	2.56	.00029
Schiff (1992/1993)	25	.16 (.09)	4.97	.00004	3.24	<.00001
Waehler (1991)	28	.25 (.12)	7.91	<.00001	5.76	<.00001
Wald et al. (1990)	28	.23 (.12)	7.03	<.00001	5.76	<.00001
Zacker (1997)	53	.20 (.10)	9.38	<.00001	4.00	<.00001
Total (except norms)	608	.19 (.11)	24.77	<.00001	4.84	<.00001
Afr						
Norms	700	.69 (.16)				
DeLucas (1997)	30	.42 (.17)	−8.54	<.00001	1.13	.29383
Erstad (1995/1996)	12	.51 (.25)	−2.48	.03001	2.44	.00542
Erstad (1995/1996)	33	.48 (.14)	−8.36	<.00001	.76	.82215
Frueh & Kinder (1994)	20	.41 (.12)	−10.18	<.00001	1.78	.06757
Goldfinger (1998/1999)	21	.47 (.17)	−5.85	<.00001	1.13	.31371
Hayslip et al. (1992)	32	.53 (.22)	−4.06	.00029	1.89	.00267
Jansak(1996/1997)	30	.46 (.19)	−6.53	<.00001	1.41	.07603
Kadle(1989)	30	.53 (.19)	−4.54	<.00001	1.41	.07603
Lipkin (1988/1989)	15	.58 (.21)	−2.02	.06290	1.72	.04694
Meisner (1988)	29	.51	−5.94	<.00001	—	—
Meyer (1989/1991)	265	.47 (.18)	−17.45	<.00001	1.26	.00916
Paul (1987/1989)	60	.53 (.19)	−6.33	<.00001	1.41	.02658
Schiff (1992/1993)	25	.58 (.19)	−2.86	.00841	1.41	.09267
Van Horn (1996)	30	.58 (.33)	−1.82	.07942	4.25	<.00001
Waehler (1991)	28	.44 (.16)	−8.11	<.00001	1.00	.46583
Zacker (1997)	53	.43 (.13)	−13.79	<.00001	1.51	.03108
Zlotogorski et al. (1987)	32	.56	−4.49	.00008	—	—
Total (except norms)	745	.49 (.19)	−21.69	<.00001	1.41	<.00001
FC						
Norms	700	4.09 (1.88)				
DeLucas (1997)	30	0.97 (0.84)	−18.46	<.00001	5.01	<.00001
Erstad (1995/1996)	12	2.42 (2.31)	−2.49	.02959	1.51	.12295
Erstad (1995/1996)	33	0.91 (1.16)	−14.86	<.00001	2.63	.00068
Goldfinger (1998/1999)	21	1.62 (1.40)	−7.87	<.00001	1.80	.05763
Greenwald (1990)	62	3.03 (1.94)	−4.13	.00010	1.06	.34895
Hallett (1996)	126	1.26 (1.33)	−20.48	<.00001	2.00	<.00001
Hayslip et al. (1992)	32	2.25 (1.72)	−5.89	<.00001	1.19	.27881
Kadle(1989)	30	3.07 (3.08)	−1.80	.08202	2.68	.00001
Kranau (1983/1984)	60	2.22	−7.39	<.00001	—	—
Meyer (1989/1991)	265	1.82 (1.76)	−17.54	<.00001	1.14	.10346
Perry & Kinder (1992)	20	1.50 (1.47)	−7.70	<.00001	1.64	.10038

APPENDIX A: *Continued*

Variable/Study	*N*	Mean (SD)	*t*	*p*	*F*	*p*
Paul (1987/1989)	60	1.37 (1.45)	−13.58	<.00001	1.68	.00687
Schiff (1992/1993)	25	3.24 (1.79)	−2.33	.02793	1.10	.40645
Sloan et al. (1996)	25	1.40 (1.40)	−9.31	<.00001	1.80	.04076
Van Horn (1996)	30	2.20 (1.90)	−5.34	.00001	1.02	.43589
Zacker (1997)	53	0.93 (1.09)	−19.07	<.00001	2.97	<.00001
Zlotogorski et al. (1987)	32	2.50	−4.68	.00004	—	—
Total (except norms)	916	1.82 (1.67)	−25.23	<.00001	1.27	.00055
Populars						
Norms	700	6.89 (1.39)				
DeLucas (1997)	30	5.81 (2.07)	−2.83	.00820	2.22	.00028
Erstad (1995/1996)	12	5.92 (2.19)	−1.53	.15410	2.48	.00466
Erstad (1995/1996)	33	6.00 (2.00)	−2.53	.01638	2.07	.00055
Greenwald (1990)	62	5.48 (1.66)	−6.49	<.00001	1.43	.02133
Kadle (1989)	30	6.20 (1.85)	−2.02	.05242	1.77	.00794
Kranau (1983/1984)	60	5.42	−7.86	<.00001	—	—
Meyer (1989/1991)	265	5.38 (1.76)	−12.56	<.00001	1.60	<.00000
Paul (1987/1989)	60	5.53 (2.14)	−4.84	.00001	2.37	<.00000
Perry & Kinder (1992)	20	4.95 (1.60)	−5.36	.00003	1.32	.15952
Ritzler & Nalesnik (1990)	30	6.5	−1.50	.14231	—	—
Schiff (1992/1993)	25	6.28 (2.17)	−1.40	.17533	2.44	.00016
Zacker (1997)	53	4.66 (1.93)	−8.25	<.00001	1.93	.00015
Zlotogorski et al. (1987)	32	5.16	−6.88	<.00001	—	—
Total (except norms)	712	5.50 (1.87)	−15.87	<.00001	1.81	<.00001
SumY						
Norms	700	0.57 (1.00)				
DeLucas (1997)	30	1.52 (1.41)	3.65	.00098	1.99	.00167
Erstad (1995/1996)	12	1.42 (1.68)	1.75	.10806	2.82	.00129
Erstad (1995/1996)	33	0.55 (0.75)	−.15	.88380	1.78	.02354
Frueh & Kinder (1994)	20	1.15 (1.84)	1.40	.17625	3.38	<.00001
Goldfinger (1998/1999)	21	1.52 (2.14)	2.03	.05595	4.58	<.00001
Greenwald (1990)	62	1.77 (1.64)	5.67	<.00001	2.69	<.00001
Kadle (1989)	30	1.27 (1.39)	2.73	.01050	1.93	.00253
Meisner (1988)	29	0.79	1.16	.25471	—	—
Meyer (1989/1991)	265	1.99 (2.20)	10.12	<.00001	4.84	<.00001
Paul (1987/1989)	60	0.73 (0.92)	1.28	.20339	1.18	.21344
Ritzler & Nalesnik (1990)	30	1.10	2.84	.00778	—	—
Schiff (1992/1993)	25	5.64 (4.93)	5.14	.00003	24.30	<.00001
Zlotogorski et al. (1987)	32	2.47	10.51	<.00001	—	—
Total (except norms)	649	1.74 (2.09)	12.95	<.00001	4.37	<.00001
SumT						
Norms	700	1.03 (0.58)				
Burns (1993/1994)	70	1.40 (1.23)	2.49	.01511	4.50	<.00001
DeLucas (1997)	30	0.42 (0.67)	−4.91	.00003	1.33	.11398
Erstad (1995/1996)	12	1.08 (1.68)	.10	.91981	8.39	<.00001
Erstad (1995/1996)	33	1.27 (1.42)	.97	.34063	5.99	<.00001
Frueh & Kinder (1994)	20	0.20 (0.52)	−7.01	<.00001	.80	.70395
Greenwald (1990)	62	0.63 (1.02)	−3.04	.00337	3.09	<.00001
Kadle(1989)	30	0.93 (1.23)	−.44	.66085	4.50	<.00001
Meisner (1988)	29	0.69	−3.09	.00422	—	—
Meyer (1989/1991)	265	0.65 (0.91)	−6.33	<.00001	2.46	<.00001
Paul (1987/1989)	60	1.13 (1.24)	.62	.53828	4.57	<.00001
Perry & Kinder (1992)	20	0.60 (0.82)	−2.33	.03074	2.00	.00701
Ritzler & Nalesnik (1990)	30	1.30	2.50	.01795	—	—
Schiff (1992/1993)	25	3.56 (2.04)	6.19	<.00001	12.37	<.00001
Waehler (1991)	28	0.75 (1.04)	−1.42	.16797	3.22	<.00001
Zacker (1997)	53	0.36 (0.59)	−7.98	<.00001	1.03	.41013
Zlotogorski et al. (1987)	32	0.66	−3.53	.00122	—	—
Total (except norms)	799	0.87 (1.07)	−3.66	.00026	3.40	<.00001
WSumC						
Norms	700	4.52 (1.79)				
DeLucas (1997)	30	1.57 (1.17)	−13.16	<.00001	2.34	.00332
Erstad (1995/1996)	12	4.08 (2.84)	−.53	.60324	2.52	.00410
Erstad (1995/1996)	33	2.20 (1.74)	−7.48	<.00001	1.06	.44379
Goldfinger (1998/1999)	21	3.09	−3.61	.00163	—	—
Hallett (1996)	126	3.21	−7.56	<.00001	—	—
Hayslip et al. (1992)	32	3.02 (2.04)	−4.09	.00026	1.30	.13012
Kadle(1989)	30	4.57 (3.33)	.08	.93540	3.46	<.00001
Meyer (1989/1991)	265	2.80 (2.00)	−12.26	<.00001	1.25	.01318
Paul (1987/1989)	60	2.67 (2.29)	−6.10	<.00001	1.64	.00251
Ritzler & Nalesnik (1990)	30	4.40	−.36	.72157	—	—
Schiff (1992/1993)	25	6.40 (3.37)	2.77	.01040	3.54	<.00001

APPENDIX A: *Continued*

Variable/Study	N	Mean (SD)	t	p	F	p
Van Horn (1996)	30	3.70	−2.46	.01970	—	—
Zacker (1997)	53	1.99 (1.99)	−8.98	<.00001	1.24	.12896
Zlotogorski et al. (1987)	32	3.07	−4.48	.00008	—	—
Total (except norms)	779	3.06 (2.18)	−14.13	<.00001	1.48	<.0001
MOR						
Norms	700	0.70 (0.82)				
Alexander (1997/1998)	52	0.81	.93	.35448	—	—
Burns (1993)	70	0.89 (1.19)	1.30	.19577	2.11	<.00001
Caine et al. (1995)	20	0.40 (0.50)	−2.58	.01686	2.69	<.00642
DeLucas (1997)	30	1.03 (1.08)	1.65	.10855	1.73	.01020
Erstad (1995/1996)	12	1.33 (1.15)	1.89	.08499	1.97	.02905
Erstad (1995/1996)	33	0.94 (1.34)	1.02	.31516	2.67	<.00001
Frueh & Kinder (1994)	20	0.70 (1.08)	.00	1.00000	1.73	.02676
Goldfinger (1998/1999)	21	2.62 (2.75)	3.20	.00452	11.25	<.00001
Greenwald (1990)	62	1.50 (1.66)	3.75	.00038	4.10	<.00001
Hallett (1996)	126	0.79 (1.03)	.93	.35421	1.58	.00021
Hayslip et al. (1992)	32	0.75 (0.92)	.30	.76453	1.26	.15965
Hilsenroth (1996/1997)	50	1.10	3.33	.00153	—	—
Kadle(1989)	30	0.43 (0.68)	−2.11	.04259	1.45	.10836
Meisner (1988)	29	0.72	.13	.89844	—	—
Meyer (1989/1991)	265	1.43 (1.51)	7.46	<.00001	3.39	<.00001
Paul (1987/1989)	60	0.67 (1.05)	−.22	.82986	1.64	.00242
Ritzler & Nalesnik (1990)	30	0.90	1.31	.20027	—	—
Schiff (1992/1993)	25	1.20 (1.04)	2.38	.02536	1.61	.03335
Zlotogorski et al. (1987)	32	0.84	−.94	.35162	—	—
Total (except norms)	999	1.07 (1.34)	7.04	<.00001	2.67	<.00001
WSum6						
Norms	700	3.28 (2.89)				
Burns (1993/1994)	70	5.41 (7.15)	2.47	.01581	6.12	<.00001
DeLucas (1997)	30	4.03 (4.55)	.90	.37784	2.48	.00003
Erstad (1995/1996)	12	6.33 (6.87)	1.54	.15272	5.65	<.00001
Erstad (1995/1996)	33	7.09 (8.26)	2.64	.01259	8.17	<.00001
Frueh & Kinder (1994)	20	9.70 (11.39)	2.52	.02088	15.53	<.00001
Greenwald (1990)	62	2.77 (3.76)	−1.04	.30153	1.69	.00114
Netter & Viglione (1994)	20	9.00 (12.34)	2.07	.05214	18.23	<.00001
Paul (1987/1989)	60	1.90 (3.42)	−3.03	.00344	1.40	.02912
Perry & Kinder (1992)	20	1.55 (2.60)	−2.92	.00828	1.24	.30286
Ritzler & Nalesnik (1990)	30	3.80	.96	.34186	—	—
Schiff (1992/1993)	25	17.92 (16.69)	4.38	.00020	33.35	<.00001
Total (except norms)	382	5.40 (7.78)	5.14	<.00001	7.24	<.00001
Lambda						
Norms	700	0.58 (0.26)				
DeLucas (1997)	30	1.37 (1.56)	2.77	.00962	36.00	<.00001
Erstad (1995/1993)	12	1.17 (2.47)	.83	.42560	90.25	<.00001
Erstad (1995/1993)	33	1.29 (2.19)	1.86	.07183	70.95	<.00001
Frueh & Kinder	20	0.68 (0.58)	.77	.45133	4.98	<.00001
Greenwald (1990)	62	0.83 (0.57)	3.42	.00109	4.81	<.00001
Hayslip et al. (1992)	32	0.62 (0.36)	.62	.53882	1.92	.00218
Jansak (1996/1997)	30	0.94 (0.83)	2.37	.02456	10.19	<.00001
Kadle (1989)	30	0.83 (1.11)	1.23	.22774	18.23	<.00001
Meyer (1989/1991)	265	0.58 (0.16)	.00	1.00000	2.64	<.00001
Paul (1987/1989)	60	0.97 (0.83)	3.62	.00060	10.19	<.00001
Schiff (1992/1993)	25	0.62 (0.46)	.43	.66928	3.13	<.00001
Sloan et al. (1996)	25	1.00 (0.80)	2.62	.01496	9.47	<.00001
Van Horn (1996)	30	0.69 (0.43)	1.39	.17470	2.74	<.00001
Waehler (1991)	28	1.26 (1.45)	2.48	.01965	31.10	<.00001
Zacker (1997)	53	2.02 (4.06)	2.58	.01269	243.84	<.00001
Zlotogorski et al. (1987)	32	0.66	1.70	.09789	—	—
Total (except norms)	767	0.88 (1.38)	5.91	<.00001	28.17	<.00001
Pure H						
Norms	700	3.40 (1.80)				
Burns (1993/1994)	70	2.76 (1.56)	−3.22	.00177	1.33	.06787
DeLucas (1997)	30	1.90 (1.04)	−7.44	<.00001	3.00	.00033
Erstad (1995/1996)	12	2.25 (1.60)	−2.46	.03067	1.26	.34982
Erstad (1995/1996)	33	1.85 (1.84)	−4.73	.00004	1.04	.40106
Frueh & Kinder	20	2.55 (1.79)	−2.09	.04917	1.01	.52805
Paul (1987/1989)	60	1.52 (1.30)	−10.38	<.00001	1.92	.00116
Schiff (1992/1993)	25	2.84 (2.07)	−1.33	.19384	1.32	.13920
Zacker (1997)	53	1.70 (1.69)	−7.03	<.00001	1.13	.29117
Total (except norms)	303	2.12 (1.59)	−11.24	<.00001	1.28	.00639

APPENDIX B: PERCENTAGE OF PROTOCOLS THAT ARE AMBITENT: NORMS (EXNER, 1993) AS COMPARED WITH NINE INDEPENDENT SAMPLES OF NONPATIENT ADULTS

Study	*N*	Ambitent Protocols		*z*	*p*
		No.	%		
Norms	700	143	20		
Goldfinger (1998/1999)	21	6	29	.91	.36385
Jansak (1996/1997)	30	14	47	3.42	.00061
Kadle (1989)	30	19	63	5.54	<.00001
Paul (1987/1989)	60	32	53	5.81	<.00001
Schiff (1992/1993)	25	13	52	3.77	.00016
Smith et al. (1991)	15	1	7	−1.32	.18851
Waehler (1991)	28	12	43	2.84	.00447
Wald et al. (1990)	28	15	54	4.17	.00003
Zacker (1997)	53	26	50	4.82	<.00001
Total (except norms)	290	138	48	8.62	<.00001

APPENDIX C: PERCENTAGE OF PROTOCOLS WITH AT LEAST ONE REFLECTION RESPONSE: NORMS (EXNER, 1993) AS COMPARED WITH EIGHT INDEPENDENT SAMPLES OF NONPATIENT ADULTS

	N	Protocols With at Least One Reflection		*z*	*p*
		No.	%		
Norms	700	47	7		
Calkins (1980/1981)	72	28	39	8.78	<.00001
DeLucas (1997)	30	9	30	4.69	<.00001
Erstad (1995/1996)	12	6	50	5.66	<.00001
Erstad (1995/1996)	33	5	15	1.84	.06505
Jacques (1990/1991)	106	38	36	9.10	<.00001
Jansak (1996/1997)	30	7	23	3.40	.00066
Paul (1987/1989)	60	6	10	.96	.33756
Schiff (1992/1993)	25	7	28	3.98	.00007
Total (except norms)	368	106	29	9.79	<.00001

NOTES

1. We acknowledge the members of the internet Rorschach Discussion list, including Michael Bridges, Jack Gerber, Rick Kramer, Helge Malmgren, Gregory Meyer, and Barry Ritzler, whose comments suggested some of the explanations discussed in this section.

2. In a recent message to the internet Rorschach Discussion Group, Gregory Meyer (personal communication, April 7, 2000) has expressed doubts that the scoring and administration of Rorschachs for his dissertation (Meyer, 1989/1991) was adequate. However, his retrospective reevaluation does not appear to be based on a reanalysis of the original data. The dissertation states: "I blindly scored 30 randomly chosen protocols from the data pool. . . . Exact agreement was found for 87.5% of the scores given. . . . The scoring reliability increased to 96.1% exact agreement when agreements were determined by score inclusion *and* score exclusion. This was in line with the interscorer agreements reported by Exner (1986) and compared favorably with the estimates of reliability found by other investigators using the Exner system" (p. 157). The dissertation futher states: "In summary, there were no clear problems with the present sample in terms of Rorschach scoring or in terms of its comparability to a typical college student population" (p. 175).

3. The most recent edition of *TRACS* (Exner, 1993, pp. 433–434) underreports the frequency of Reflection responses among nonpatients in the 1969 study. Specifically, although the original study (Exner, 1969) reported a rate of 35% (7 protocols with Reflection responses), the recent edition of *TRACS* (Exner, 1993) underreports the rate as 15% (3 protocols with Reflections responses).

4. For this particular set of comparisons, alpha was set liberally at .05. If alpha were to be adjusted using the Bonferroni correction (.05/15 = .0033), none of the changes over time would be statistically significant.

5. Each of these studies, except Loucks et al. (1980), included at least one nonabused group of children that fit the definition for nonpatients as described in the search criteria for the present review.

REFERENCES

Ackerman, M. J., & Ackerman, M. C. (1997). Custody evaluation practices: A survey of experienced professionals (revisited). *Professional Psychology: Research and Practice, 28,* 137–145.

Acklin, M. (1990). Personality dimensions in two types of learning-disabled children: A Rorschach study. *Journal of Personality Assessment, 54,* 67–77.

Acklin, M. W. (1999). Behavioral science foundations of the Rorschach test: Research and clinical applications. *Assessment, 6,* 319–326.

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.

Alexander, S. (1998). A comparison of black Americans from single-and two-parent families on the Rorschach (Doctoral dissertation, California School of Professional Psychology, Berkeley/Alameda, 1997). *Dissertation Abstracts International, 58,* 5184B.

American Psychological Association (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47,* 1597–1611.

American Psychological Association (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.

Archer, R. P. (1999). Introduction to a special section: Perspectives on the Rorschach. *Assessment, 6,* 307–311.

Aronow, E. (1999). The Rorschach: An integrative approach. *Contemporary Psychology, 44,* 546–547

Aronow, E., Reznikoff, M., & Moreland, K. (1994). *The Rorschach technique: Perceptual basics, content interpretation, and applications*. Boston: Allyn and Bacon.

Aronow, E., Reznikoff, M., & Moreland, K. L. (1995). The Rorschach: Projective technique or psychometric test? *Journal of Personality Assessment, 64,* 213–228.

Board of Professional Affairs (1998). Awards for distinguished professional contributions: John Exner. *American Psychologist, 53,* 391–392.

Bridges, M. R., Wilson, J. S., & Gacono, G. B. (1998). A Rorschach investigation of defensiveness, self-perception, interpersonal relations, and affective states in incarcerated pedophiles. *Journal of Personality Assessment, 70,* 365–385.

Burns, B. (1994). An object relations study: Relationship between the Rorschach Human Experience Variable and interpersonal relatedness among nonpatients (Doctoral dissertation, California School of Professional Psychology, San Diego, 1993). *Dissertation Abstracts International, 54,* 3847B.

Burns, B., & Viglione, D. J. (1996). The Rorschach Human Experience variable, interpersonal relatedness, and object representation in nonpatients. *Psychological Assessment, 8,* 92–99.

Burns, B., & Viglione, D. J. (1997). Correction to Burns and Viglione (1996). *Psychological Assessment, 9,* 82.

Caine, S. L., Frueh, B. C., & Kinder, B. N. (1995). Rorschach susceptibility to malingered depressive disorders in adult females. In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in Personality Assessment, Volume 10* (pp. 165–174). Hillsdale, NJ: Lawrence Erlbaum.

Calkins, D. L. (1981). The Rorschach Egocentricity Index: A study of its relationship to self esteem (Doctoral dissertation, Brigham Young University, 1980). *Dissertation Abstracts International, 42,* 365B.

Celenza, A., & Hilsenroth, M. (1997). Personality characteristics of mental health professionals who have engaged in sexualized dual relationships: A Rorschach investigation. *Bulletin of the Menninger Clinic, 61,* 90–107.

Champion, L., Doughtie, E. B., Johnson, P. J., & McCreary, J. H. (1984). Preliminary investigation into the Rorschach response patterns of children with documented learning disabilities. *Journal of Clinical Psychology, 40,* 329–333.

Costello, R. M. (1999). Two foundations of Rorschach assessment revisited. *European Journal of Psychological Assessment, 15,* 64–77.

Cruz, E. B., Brier, N. M., & Reznikoff, M. (1997). An examination of the relationship between form level ratings on the Rorschach and learning disability status. *Journal of Psychology, 131,* 167–174.

Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on myth*. New York: Free Press.

DeLucas, A. L. (1997). The quality of paternal attachment and violent behaviors: A Rorschach study of convicted military male offenders (Doctoral dissertation, California School of Professional Psychology, San Diego, 1997). *Dissertation Abstracts International, 58,* 2115B.

DeSousa, D. T. (1993). Object relations, object representations and social functioning of sexually abused girls (Unpublished dissertation, Loyola University of Chicago, 1993). *Dissertation Abstracts International, 53,* 6545B.

Ebert, M. M. (1992). An analysis of the Rorschach as a tool to differentiate children with attention-deficit hyperactivity disorder from controls (Unpublished dissertation, Brigham Young University, 1991). *Dissertation Abstracts International, 52,* 4972B.

Erstad, D. (1996). An investigation of older adults' less frequent human movement and color responses on the Rorschach (Doctoral dissertation, Marquette University, 1995). *Dissertation Abstracts International, 57,* 4084B.

Esmail, J. J. (1997). Treatment needs of non-abused siblings of child sexual abuse victims in the PACT (Project Aimed at Coordinated Treatment) program (Unpublished dissertation, Spalding University, 1996). *Dissertation Abstracts International, 57,* 5323B.

Exner, J. E. (1969). Rorschach responses as an index of narcissism. *Journal of Projective Techniques and Personality Assessment, 33,* 324–330.

Exner, J. E. (1974). *The Rorschach: A Comprehensive System. Volume 1.* New York: Wiley.

Exner, J. E. (1978). *The Rorschach: A Comprehensive System. Volume 2. Current research and advanced interpretation.* New York: Wiley.

Exner, J. E. (1985). *A Rorschach workbook for the Comprehensive System* (2nd ed.). Asheville, NC: Rorschach Workshops.

Exner, J. E. (1986). *The Rorschach: A Comprehensive System. Volume 1: Basic Foundations* (2nd ed.). New York: Wiley.

Exner, J. E. (1990). *A Rorschach workbook for the Comprehensive System* (3rd ed.). Asheville, NC: Rorschach Workshops.

Exner, J. E. (1991). *The Rorschach: A Comprehensive System. Volume 2: Interpretation* (2nd ed.). New York: Wiley.

Exner, J. E. (1993). *The Rorschach: A Comprehensive System, Volume 1: Basic foundations* (3rd ed.). New York: Wiley.

Exner, J. E. (1995a). Comment on "Narcissism in the Comprehensive System for the Rorschach." *Clinical Psychology: Science and Practice, 2,* 200–206.

Exner, J. E. (1995b). *A Rorschach workbook for the Comprehensive System* (4th ed.). Asheville, NC: Rorschach Workshops.

Exner, J. E. (1996). A comment on "The Comprehensive System for the Rorschach: A critical examination." *Psychological Science, 7,* 11–13.

Exner, J. E., Colligan, S. C., Boll, T. J., Stischer, B., & Hillman, L. (1996). Rorschach findings concerning closed head injury patients. *Assessment, 3,* 317–326.

Exner, J. E., Kinder, B. N., & Curtiss, G. (1995). Reviewing basic design features. In J. E. Exner (Ed.), *Issues and methods in Rorschach research* (pp. 145–158). Mahwah, NJ: Lawrence Erlbaum.

Exner, J. E., & Sendin, C. (1997). Some issues in Rorschach research. *European Journal of Psychological Assessment, 13,* 155–163.

Exner, J. E., & Weiner, I. B. (1982). *The Rorschach: A Comprehensive System, Volume 3: Assessment of children and adolescents.* New York: Wiley.

Exner, J. E., & Weiner, I. B. (1995). *The Rorschach: A Comprehensive System, Volume 3: Assessment of children and adolescents* (2nd ed.). New York: Wiley.

Exner, J. E., Weiner, I. B., & Schyler, W. (1976). *A Rorschach workbook for the Comprehensive System.* Bayville, NY: Rorschach Workshops.

Exner, J. E., Weiss, L. J., Coleman, M., & Rose, R. B. (1979). Rorschach variables for a group of occupationally successful dancers. Workshops study no. 250 (unpublished), Rorschach Workshops.

Eysenck, H. J. (1959). The Rorschach Inkblot Test. In O. K. Buros (Ed.), *The fifth mental measurements yearbook* (pp. 276–278). Highland Park, NJ: Gryphon Press.

Finn, S. E. (1996). Assessment feedback integrating MMPI-2 and Rorschach findings. *Journal of Personality Assessment, 67,* 543–557.

Frueh, B. C., & Kinder, B. N. (1994). The susceptibility of the Rorschach Inkblot Test to malingering of combat-related PTSD. *Journal of Personality Assessment, 62,* 280–298.

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: Erlbaum.

Ganellen, R. J. (1994). Attempting to conceal psychological disturbance: MMPI defensive response sets and the Rorschach. *Journal of Personality Assessment, 63,* 423–437.

Ganellen, R. J. (1996a). Comparing the diagnostic efficiency of the MMPI, MCMI-II, and Rorschach: A review. *Journal of Personality Assessment, 67,* 219–243.

Ganellen, R. J. (1996b). *Integrating the Rorschach and the MMPI-2 in personality assessment.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Gann, M. K. (1995). The Rorschach and other projective methods. In Jay Ziskin (Ed.), *Coping with psychiatric and psychological testimony. Vol. II: Special topics* (5th ed.) (pp. 823–884). Los Angeles: Law and Psychology Press.

Garb, H. N. (1984). The incremental validity of information used in personality assessment. *Clinical Psychology Review, 4,* 641–655.

Garb, H. N. (1998). *Studying the clinician: judgment research and psychological assessment.* Washington, DC: American Psychological Association.

Garb, H. N. (1999). Call for a moratorium on the use of the Rorschach Inkblot in clinical and forensic settings. *Assessment, 6,* 313–315.

Garb, H. N., Wood, J. M., & Lilienfeld, S. O. (2000). The detection and assessment of child sexual abuse: An evaluation of the Rorschach, Thematic Apperception Test, and projective drawings. Manuscript in preparation.

Garb, H. N., Wood, J. M., & Nezworski, M. T. (2000). Projective techniques and the detection of child sexual abuse. *Child Maltreatment, 5,* 161–168.

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.*

Garfield, S. L. (2000). The Rorschach test in clinical diagnosis – A brief commentary. *Journal of Clinical Psychology, 56,* 431–434.

Goldfinger, D. (1999). Rorschach patterns in Vietnam veterans with post-traumatic stress disorder: A study of cognition, psychophysiology, and psychological defense (Doctoral dissertation, University of Michigan, 1998). *Dissertation Abstracts International, 59,* 3691B.

Greenwald, D. F. (1990). An external construct validity study of Rorschach personality variables. *Journal of Personality Assessment, 55,* 768–780.

Hallett, S. J. (1996). Trauma and coping in homicide and child sexual abuse detectives (Doctoral dissertation, California School of Professional Psychology, San Diego, 1996). *Dissertation Abstracts International, 57,* 2152B.

Harper, G., & Scott, R. (1990). Learning disabilities: An appraisal of Rorschach response patterns. *Psychological Reports, 67,* 691–696.

Hartman, W. L., Clark, M. E., Morgan, M. K., Dunn, V. K., Fine, A. D., Perry, G. G., & Winsch, D. L. (1990). Rorschach structure of a hospitalized sample of Vietnam veterans with PTSD. *Journal of Personality Assessment, 54,* 149–159.

Hayslip, B., McBride, P. A., Lowman, R. L., & Aronson, H. J. (1992). Color vision deficits and Rorschach performance in aged persons. *International Journal of Aging and Human Development, 34,* 165–173.

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.

Hilsenroth, M. J. (1997). A multidimensional assessment of narcissistic personality disorder (Doctoral dissertation, University of Tennessee, Knoxville, 1996). *Dissertation Abstracts International, 57,* 4709B.

Hilsenroth, M. J., Fowler, J. C., Padawer, J. R., & Handler, L. (1997). Narcissism in the Rorschach revisited: Some reflections on empirical data. *Psychological Assessment, 9,* 113–121.

Holaday, M. (1998). Rorschach protocols of children and adolescents with severe burns: A follow-up study. *Journal of Personality Assessment, 7,* 306–321.

Holaday, M., & Whittenberg, T. (1994). Rorschach responding in children and adolescents who have been severely burned. *Journal of Personality Assessment, 62,* 269–279.

Hunsley, J., & Bailey, J. M. (1999). The clinical utility of the Rorschach: Unfulfilled promises and an uncertain future. *Psychological Assessment, 11,* 266–277.

Jacques, M. F. (1991). The Rorschach Egocentricity Index: A validation study (Doctoral dissertation, California School of Professional Psychology, San Diego, 1990). *Dissertation Abstracts International, 52,* 1064B.

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, San Diego, 1996). *Dissertation Abstracts International, 57,* 6576B.

Jensen, A. R. (1965). The Rorschach Inkblot Test. In O. K. Buros (Ed.), *The sixth mental measurements yearbook* (pp. 501–509). Highland Park, NJ: Gryphon Press.

Jorgensen, K., Andersen, T. J., & Dam, H. (in press).The diagnostic efficiency of the Rorschach Depression Index and the Schizophrenia Index: A review. *Assessment.*

Kadle, J. C. (1989). The suitability of standard Rorschach scoring categories for nonpatient and clnically depressed elderly (Doctoral dissertation, Case Western Reserve University, 1989). *Dissertation Abstracts International, 50,* 2154B.

Kaser-Boyd, N. (1993). Rorschachs of women who commit homicide. *Journal of Personality Assessment, 60,* 458–470.

Kelly, R. J., & Ben-Meir, S. (1993). Emotional effects. In J. Waterman, R. J. Kelly, M. K. Oliveri, & J. McCord (Eds.), *Behind the playground walls: Sexual abuse in preschools* (pp. 106–119). New York: Guilford.

Kranau, E. J. (1984). Level of acculturation/biculturation and Rorschach protocols of Hispanic Americans and Anglo Americans (Doctoral dissertation, Oklahoma State University, 1983). *Dissertation Abstracts International, 44,* 3200B.

Kubiszyn, T. W., Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., & Eisman, E. J. (2000). Empirical support for psychological assessment in clinical health care settings. *Professional Psychology: Research and Practice, 31,* 119–130.

Lees-Haley, P. R. (1992). Psychodiagnostic test usage by forensic psychologists. *American Journal of Forensic Psychology, 10,* 25–30.

Leifer, M., Shapiro, J. P., Martone, M. W., & Kassem, L. (1991). Rorschach assessment of psychological functioning in sexually abused girls. *Journal of Personality Assessment, 56,* 14–28.

Lerner, P. M. (2000). A nonreviewer's comment: On the Rorschach and baseball. *Journal of Clinical Psychology, 56,* 439.

Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The scientific status of projective techniques. *Psychological Science in the Public Interest, 1,* 27–66.

Lipkin, L. R. (1989). Rorschach performance as a test of a theory of late-onset depression in women (Doctoral dissertation, Temple University, 1988). *Dissertation Abstracts International, 49,* 2864B.

Lipsey, M. W. (1990). *Design sensitivity: Statistical power for experimental research.* Newbury Park, CA: Sage.

Loucks, S., Burstein, A. G., Boros, T., & Kregor, E. (1980). The affective ratio in Rorschach's test as a function of age. *Journal of Personality Assessment, 44,* 590–591.

Meisner, S. (1988). Susceptibility of Rorschach distress correlates to malingering. *Journal of Personality Assessment, 52,* 564–571.

Meyer, G. J. (1991). An empirical search for fundamental personality and mood dimensions within the Rorschach

test (Unpublished 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.

Netter, B. E. C., & Viglione, D. J. (1994). An empirical study of malingering schizophrenia on the Rorschach. *Journal of Personality Assessment, 62,* 45–57.

Nezworski, M. T., & Wood, J. M. (1995). Narcissism in the Comprehensive System for the Rorschach. *Clinical Psychology: Science and Practice, 2,* 179–199.

Paul, M. E. (1989). Geriatric normative data for the Rorschach for a community-living sample (Doctoral dissertation, Ball State University, 1987). *Dissertation Abstracts International, 49,* 3453B.

Perry, G. G., & Kinder, B. N. (1992). Susceptibility of the Rorschach to malingering: A schizophrenia analogue. In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in Personality Assessment. Volume 9* (pp. 127–140). Hillsdale, NJ: Erlbaum.

Pinkerman, J. E., Haynes, J. P., & Keiser, T. (1993). Characteristics of psychological practice in juvenile court clinics. *American Journal of Forensic Psychology, 11,* 3–12.

Piotrowski, C. (1999). Assessment practices in the era of managed care: Current status and future directions. *Journal of Clinical Psychology, 55,* 787–796.

Ritzler, B. A., & Exner, J. E. (1995). Special issues in subject selection and design. In J. E. Exner (Ed.), *Issues and methods in Rorschach research* (pp. 123–143). Mahwah, NJ: Erlbaum.

Ritzler, B., & Nalesnik, D. (1990). The effect of inquiry on the Exner Comprehensive System. *Journal of Personality Assessment, 55,* 647–656.

Salyer, K. M., Holmstrom, R. W., & Noshpitz, J. D. (1991). Learning disabilities as a childhood manifestation of severe psychopathology. *American Journal of Orthopsychiatry, 61,* 230–240.

Schiff, L. M. (1993). A descriptive Rorschach study of psychoanalytic psychoanalysts (Doctoral dissertation, California School of Professional Psychology, Berkeley/Alameda, 1992). *Dissertation Abstracts International, 53,* 4968B.

Sechrest, L., & McKnight, P. (2000) Rorschach as a source of information: Psychometric white spaces. Manuscript submitted for publication.

Sechrest, L., Stickle, T. R., & Stewart, M. (1998). The role of assessment in clinical psychology. In A. Bellack, M. Hersen (Series Eds.), & C. R. Reynolds (Vol. Ed.), *Comprehensive Clinical Psychology, Vol. 4: Assessment* (pp. 1–32). New York: Pergamon.

Shaffer, L. F. (1959). The Rorschach Inkblot Test. In O. K. Buros (Ed.), *The fifth mental measurements yearbook* (pp. 285–289). Highland Park, NJ: Gryphon Press.

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.

Sloan, P., Arsenault, L., Hilsenroth, M., Handler, L., & Harvill, L. (1996). Rorschach measures of posttraumatic stress in Persian Gulf War veterans: A three-year follow-up study. *Journal of Personality Assessment, 66,* 54–64.

Sloan, P., Arsenault, L., Hilsenroth, M., Harvill, L., & Handler, L. (1995). Rorschach measures of posttraumatic stress in Persian Gulf War veterans. *Journal of Personality Assessment, 64,* 397–414.

Smith, J. E., Hillard, M. C., Walsh, R. A., Kubacki, S. R., & Morgan, C. D. (1991). Rorschach assessment of purging and nonpurging bulimics. *Journal of Personality Assessment, 56,* 277–288.

Stricker, G., & Gold, J. R. (1999). The Rorschach: Toward a nomothetically based, idiographically applicable configurational model. *Psychological Assessment, 11,* 240–250.

Swanson, G. S., Blount, J., & Bruno, R. (1990). Comprehensive System Rorschach data on Vietnam combat veterans. *Journal of Personality Assessment, 54,* 160–169.

Van Horn, P. J. (1996). A Rorschach study of posttraumatic stress disorder in Vietnam theater veterans (Doctoral dissertation, Pacific Graduate School of Psychology, 1996). *Dissertation Abstracts International, 57,* 2890B.

Viglione, D. J. (1997). Problems in Rorschach research and what to do about them. *Journal of Personality Assessment, 68,* 590–599.

Viglione, D. J. (1999). A review of recent research addressing the utility of the Rorschach. *Psychological Assessment, 11,* 251–265.

Viglione, D. J., & Exner, J. E. (1995). Formulating issues in Rorschach research. In J. E. Exner (Ed.), *Issues and methods in Rorschach research* (pp. 53–71). Mahwah, NJ: Lawrence Erlbaum.

Vincent, K. R., & Harman, M. J. (1991). The Exner Rorschach: An analysis of clinical validity. *Journal of Clinical Psychology, 47,* 596–599.

Waehler, C. A. (1991). Selected Rorschach variables of never-married men. *Journal of Clinical Psychology, 47,* 123–132.

Wald, B. K., Archer, R. P., & Winstead, B. A. (1990). Rorschach characteristics of mothers of incest victims. *Journal of Personality Assessment, 55,* 417–425.

Walpole, R. E., & Myers, R. H. (1985). *Probability and statistics for engineers and scientists* (3rd ed.). New York: Macmillan.

Watkins, C. E., Campbell, V. L., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assess-

ment by clinical psychologists. *Professional Psychology: Research and Practice, 26,* 54–60.

Weiner, I. B. (1995). Methodological considerations in Rorschach research. *Psychological Assessment, 7,* 330–337.

Weiner, I. B. (1996). Some observations on the validity of the Rorschach Inkblot Method. *Psychological Assessment, 8,* 206–213.

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.

Weiner, I. B. (1999). What the Rorschach can do for you: Incremental validity in clinical applications. *Assessment, 6,* 327–339.

Weiner, I. B. (2000). Using the Rorschach properly in practice and research. *Journal of Clinical Psychology, 56,* 435–438.

Wideman, B. G. (1998). Rorschach responses in gifted and nongifted children: A comparison study (Doctoral dissertation, Georgia State University, 1998). *Dissertation Abstracts International, 59,* 905B.

Wiggins, J. S. (1988). *Personality and prediction: Principles of personality assessment*. Malabar, FL: Krieger.

Winter, L. B., & Exner, J. E. (1973). Some psychological characteristics of some successful theatrical artists. Workshops study no. 183 (unpublished), Rorschach Workshops.

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. (2000a). 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., Lilienfeld, S. O., Garb, H. N., & Nezworski, M. T. (2000b). Limitations of the Rorschach as a diagnostic tool: A reply to Garfield (2000), Lerner (2000), and Weiner (2000). *Journal of Clinical Psychology, 56,* 441–448.

Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1996a). The Comprehensive System for the Rorschach: A critical examination. *Psychological Science, 7,* 3–10.

Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1996b). Thinking critically about the Comprehensive System for the Rorschach. A reply to Exner. *Psychological Science, 7,* 14–17.

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.

Yanovsky, A. (1995). The role of denial in adolescent sex offenders who were sexually victimized as children as measured by the Rorschach (Unpublished dissertation, California School of Professional Psychology, Fresno, 1994). *Dissertation Abstracts International, 56,* 2893B.

Zacker, J. (1997). Rorschach responses of police applicants. *Psychological Reports, 80,* 523–528.

Zimmerman, D. P., & Dillard, J. (1994). The Rorschach assessment of sexually abused children in residential treatment. *Applied Research in Residential Treatment, 12,* 59–72.

Zlotogorski, Z., Hahnemann, L. E., & Wiggs, E. A. (1987) Personality characteristics of hypnotizability. *American Journal of Clinical Hypnosis, 30,* 51–56.

Received March 10, 2000; revised August 22, 2000; accepted August 30, 2000.