

# THE RORSCHACH INKBLOT TEST: A CASE OF OVERSTATEMENT?

James M. Wood  
University of Texas at El Paso

Scott O. Lilienfeld  
Emory University

In the 1940s, inflated claims were often made regarding the Rorschach Inkblot Test. Over half a century later, overstatements regarding the test are still common. The present article identifies problems with the Rorschach regarding norms, cultural sensitivity, interrater reliability, test-retest reliability, validity, factor structure, and accessibility of supporting studies. Contrary to overstated claims made on behalf of the Rorschach, the test continues to be a highly problematic instrument from a psychometric standpoint.

*Keywords:* Rorschach Inkblot Test, projective tests, psychometrics, reliability, validity

In the 1940s, the Rorschach Inkblot Test was called an x-ray of the mind, allowing psychologists to peer into the hidden recesses of the unconscious. This claim is no longer taken seriously. However, in the present day new and perhaps equally expansive claims have appeared. Two years ago, the Award for Distinguished Professional Contributions to Knowledge was bestowed on John Exner by the Board of Professional Affairs of the American Psychological Association (APA). The Board's commendation appeared in the April issue of *American Psychologist* (Board of Professional Affairs, 1998, p. 392). It stated:

---

This article is based in part on a symposium (The Rorschach: A Critical Look) presented at the 1999 meeting of the American Psychological Association in Boston.

The authors gratefully acknowledge the important contributions made by collaborators on earlier Rorschach articles: M. Teresa Nezworski, William J. Stejskal, Howard N. Garb, and William M. Grove.

Correspondence concerning this article and requests for offprints should be addressed to James M. Wood, PhD, Department of Psychology, University of Texas at El Paso, El Paso, TX 79968. E-mail: [jwood@utep.edu](mailto:jwood@utep.edu)

Exner has almost single-handedly rescued the Rorschach and brought it back to life. *The result is the resurrection of perhaps the single most powerful psychometric instrument ever envisioned.* [emphasis added]

As may be seen, the Board of Professional Affairs expressed its high esteem in language reminiscent of the New Testament: Exner was said to have brought the Rorschach "back to life," giving the test a new "resurrection." However, the present article focuses not on the Board's quasi-religious language, but instead on the second part of the quotation. Is it true that the Rorschach is "the single most powerful psychometric instrument ever envisioned?" To answer this question, we need to review the basic psychometric properties of the test: Norms, interrater reliability, test-retest reliability, and validity.

## Norms

Norms developed for the Comprehensive System (Exner, 1974, 1978, 1986, 1991, 1993; Exner & Weiner, 1982, 1995) have represented a step forward

for the Rorschach. However, do these norms justify the claim that the Rorschach is the “most powerful psychometric instrument ever envisioned?” A few facts put the situation into perspective. First, the adult norms for the Comprehensive System are based on a sample of 700 participants. By comparison, the norms for the WAIS-III (Psychological Corporation, 1997) are based on over 2,400 participants. The Comprehensive System normative data were collected in the 1970s and 1980s, the WAIS-III data in the 1990s. There does not appear to be any superiority here for the Comprehensive System.

Second, the children’s norms for the Comprehensive System are problematic. Like the adult normative data, the children’s normative data were apparently gathered 15 to 25 years ago. In the mid-1980s Exner concluded that Rorschach protocols with fewer than 14 responses were invalid. Such protocols were dropped from the Comprehensive System normative samples, including a disproportionate number of children’s protocols. As a result, normative samples for some age groups of children shrank substantially, and the stratification patterns for geographic regions and socioeconomic status were altered (Exner, 1993). Thus, the Comprehensive System normative data reported for children in recent books (Exner, 1993; Exner & Weiner, 1995) may be out of date and unrepresentative of the general population of American children.

Third, the cultural sensitivity of Rorschach norms is a cause for concern. The Comprehensive System and other Rorschach systems have virtually no normative data for American minorities (Garb, Wood, Nezworski, Grove, & Stejskal, 1999). As Bernadette Gray-Little (1995, p. 150) has observed, “In the use of inkblots with ethnic minorities, the assessor must be aware that there are few empirical data to provide a guide.”

The lack of cross-cultural norms is important because American minorities and non-Americans do score differently on the Rorschach. Studies have shown that Blacks, Hispanics, Native Americans, and non-Americans score differently on important Rorschach variables for both the Comprehensive System and other approaches

(Aposhian, 1994/1995; Baca, 1993/1994; Day, Boyer, & Devos, 1989; Devos, 1989; Devos & Miner, 1989; Glass, Bieber & Tkachuk, 1996; Krall et al., 1983; Moon & Cundick, 1983; Munson, 1992/1993; Sanchez, 1992/1993; Sangro, 1997). Furthermore, there have been a substantial number of critiques regarding cross-cultural use of the Rorschach and particularly the lack of appropriate normative data (Constantino, Flanagan, & Malgady, 1995; Cuellar, 1998; Dana, 1993, 1998; Frank, 1992, 1993; Garb, Wood, Nezworski et al., 1999; Gray-Little, 1995; Gray-Little & Kaplan, 1998; Howes & DeBlasie, 1989; Okazaki & Sue, 1995; Velasquez, 1995; Velasquez & Callahan, 1992; but also see Butcher, Nezami, & Exner, 1998; Ritzler, 1996).

Of course, the WAIS-III and other standard intelligence tests also show consistent ethnic differences, although the meaning and interpretation of these differences, particularly the approximately one standard deviation difference between Blacks and Whites, remains highly controversial. Unlike standard cognitive ability tests, however, which have repeatedly been found to exhibit little or no slope bias between Blacks and Whites (e.g., Schmidt, 1988), Rorschach scores have not been systematically examined for evidence of slope bias—in other words, differential validity—between ethnic groups.

Because there are important cross-cultural differences, and because appropriate norms have not been developed, it is doubtful whether the Comprehensive System should currently be used to evaluate members of American minority groups (see guidelines of the American Psychological Association, 1993). As Richard Dana (1993, p. 160) has said, “The Rorschach and the Exner Comprehensive versions are not recommended for routine cross-cultural applications.”

### **Interrater Reliability**

If claims regarding the psychometric superiority of the Rorschach are not based on the normative data, then are they instead based on the test’s superior interrater reliability? Until recently, it was often claimed that the Comprehensive System for

the Rorschach could be scored as reliably as the Wechsler Adult Intelligence Scale and its progeny. Exner (1978, p. 14) had said, "in two new scorer reliability studies...the .85 level of reliability was achieved or surpassed for all scorers." Similarly, Groth-Marnat (1990, p. 279) reported that scoring categories were included in the Comprehensive System only if they achieved a minimum interrater reliability of .85.

Figure 1 shows a comparison of interrater reliability numbers for the WAIS-III subtests (Psychological Corporation, 1997) and for Rorschach variables as reported by Exner in 1993. As can be seen on the left side of the figure, the median reliability of WAIS-III subtests (marked by an oval) is in the upper .90s, the maximum is close to 1.0, and the minimum is .91. The numbers for the Rorschach, as reported by Exner (1993) and

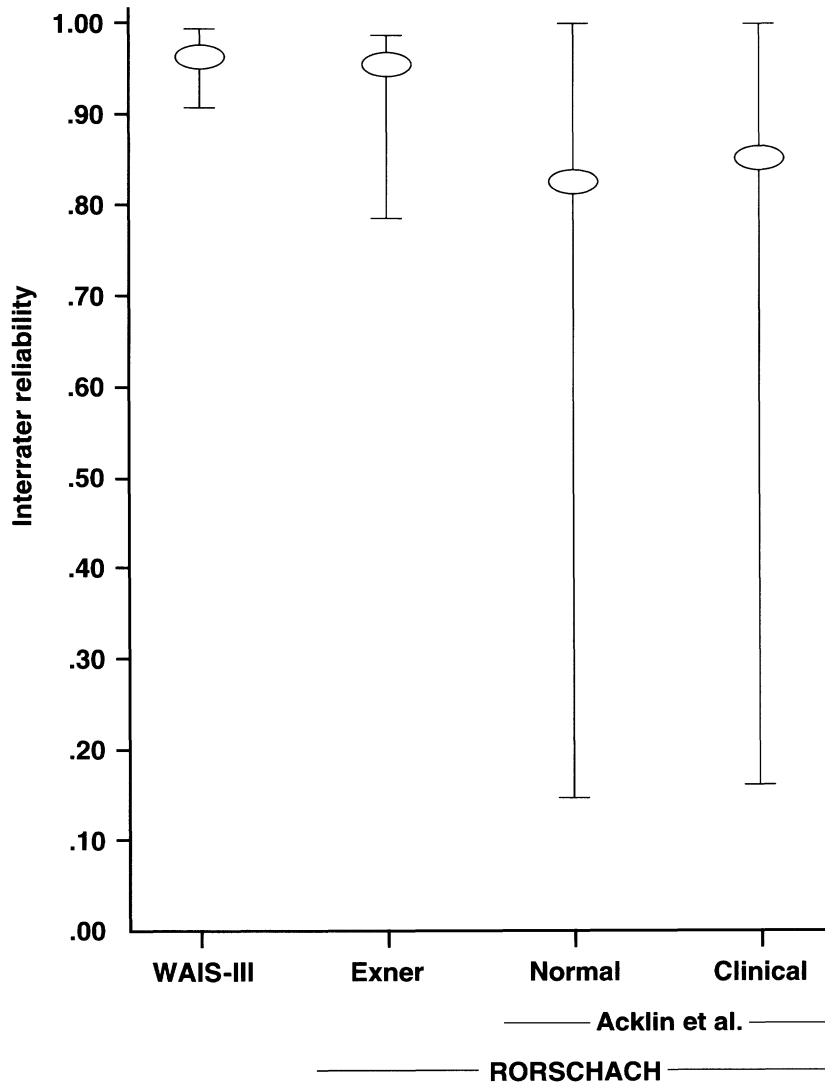


Figure 1. Maximum, median, and minimum interrater reliability for the WAIS-III (Psychological Corporation, 1997) and the Rorschach (Exner, 1993; Acklin, McDowell, & Verschell, in press).

shown second from the left in the figure, look almost as good. The median is .96, the maximum .99, and the minimum .78.

However, articles by McDowell and Acklin (1996) and Wood, Nezworski, and Stejskal (1996a, 1996b, 1997; but see Exner, 1996; Meyer, 1997a, 1997c) have pointed out that the numbers reported by Exner (1993) are for percentage agreement, a statistic that often yields inflated values because of its failure to correct for base rate agreement. Because of the shortcomings of percentage agreement as a reliability index, virtually all statistical experts today recommend the use of intraclass correlation coefficients or kappa for the assessment of interrater reliability. For example, the interrater reliability numbers shown in Figure 1 for the WAIS-III are intraclass correlation coefficients, not percentage agreement.

Recently Acklin, McDowell, and Verschell (in press) have computed the intraclass correlation coefficients for about 90% of Comprehensive System scores in two samples, as shown on the right side of Figure 1. The results for both samples of Acklin and his colleagues are very similar: The median reliability of Comprehensive System scores is in the low .80s, the maximum is 1.0, and the minimum is below .20. As Acklin and his colleagues point out, the reliability of many Comprehensive System scores is acceptable or even excellent. However, about 50% of the reliability figures for the Rorschach fall below the minimum .85 level set forth by Exner (1978) and Groth-Marnat (1990). Furthermore, several important Comprehensive System variables exhibit a level of reliability that is very problematic for test scores used in clinical work. For example, the interrater reliabilities of the Schizophrenia Index in Acklin's two samples were .452 and .560. Similar low reliability was exhibited for Adjusted *D* (.533 and .678), *X-%* (.621 and .656), and *FC:CF + C* (.543 and .165). The unfavorable numbers raise serious doubt whether these scores should be used in clinical or forensic contexts. As may be seen, the interrater reliability of the Rorschach is clearly not equal to the WAIS and its progeny, despite optimistic claims that the two have equivalent

reliability. And of course, interrater reliability is really not even an issue for MMPI scales, which can be scored by a secretary or computer.

### Test-Retest Reliability

It has often been claimed that the test-retest reliability of Comprehensive System variables is very high. For example, Meyer (1997c, p. 496) has stated "With the exception of variables thought to assess transient emotional reactions, the temporal stability of CS scores is excellent." Similarly, Viglione (in press) has claimed that "...the great majority of Rorschach Comprehensive System (CS) variables and configurations have shown impressive temporal consistency reliability [sic]..."

There are 12 books and articles by Exner and his colleagues that have reported test-retest reliability coefficients, and these coefficients are generally quite good (Exner 1974, 1978, 1980, 1986, 1988, 1991, 1993; Exner, Armbruster, & Viglione, 1978; Exner, Thomas, & Mason, 1985; Exner & Weiner, 1982, 1995; Haller & Exner, 1985). However, when we reviewed all 12 articles and books, we found that they simply report the test-retest reliability for the same set of 40 variables again and again. Table 1 shows the 40 variables whose test-retest coefficients we found in Exner's publications. There are 125 or more variables in the Comprehensive System. Where are the test-retest coefficients for the remaining 85 variables? We cannot find them anywhere in Exner's books or articles. For example, Table 2 lists a sample of 40 Comprehensive System variables whose test-test reliability coefficients seem to be unreported in Exner's books and articles. As may be seen, there seem to be no test-retest figures reported for the Schizophrenia Index (*SCZI*), the Depression Index (*DEPI*), the Coping Deficit Index (*CDI*), the Hypervigilance Index (*HVI*), or Reflection responses, even though these five variables have all been designated by Exner (1991, pp. 144-145) as "Key variables" in the Comprehensive System.

It appears, therefore, that claims regarding the "psychometric superiority" of the Rorschach cannot be based on test-retest coefficients, because

Table 1

*The 40 Comprehensive System Variables for Which Test-Retest Reliability Coefficients Have Been Reported in Exner's Publications*

1. <i>R</i>	11. <i>p</i>	21. <i>Y</i>	31. <i>EA</i>
2. <i>P</i>	12. <i>C</i>	22. <i>V</i>	32. <i>ep</i>
3. <i>Zf</i>	13. <i>C + Cn</i>	23. <i>Sum Shad</i>	33. <i>es</i>
4. <i>F</i>	14. <i>FC</i>	24. <i>Col/Sh Bl</i>	34. <i>D</i>
5. <i>M</i>	15. <i>CF</i>	25. <i>FD</i>	35. <i>Adj D</i>
6. <i>M-</i>	16. <i>CF + C</i>	26. <i>Lambda</i>	36. <i>S</i>
7. <i>FM</i>	17. <i>CF + C + Cn</i>	27. <i>X+%</i>	37. <i>MOR</i>
8. <i>m</i>	18. <i>Sum C</i>	28. <i>X-%</i>	38. <i>Zd</i>
9. <i>FM + m</i>	19. <i>C'</i>	29. <i>Afr</i>	39. <i>Sum5 Spec Scores</i>
10. <i>a</i>	20. <i>T</i>	30. $3r + (2)/R$	40. <i>Wgt Sum5 Spec Scores</i>

Table 2

*A Sample of 40 Comprehensive System Variables Whose Test-Retest Reliability Coefficients Seem to be Unreported in Exner's Publications*

1. <i>SCZI</i>	11. <i>H: Other H</i>	21. <i>EB Per</i>	31. <i>INC1</i>
2. <i>DEPI</i>	12. <i>FC:CF + C</i>	22. <i>Intellect</i>	32. <i>INC2</i>
3. <i>CDI</i>	13. <i>Xu%</i>	23. <i>ISO-R</i>	33. <i>ALOG</i>
4. <i>HVI</i>	14. <i>S-%</i>	24. <i>Blends:R</i>	34. <i>CONTAM</i>
5. <i>Reflections</i>	15. <i>F+%</i>	25. <i>W:M</i>	35. <i>DV+</i>
6. <i>WSum6</i>	16. <i>F-%</i>	26. <i>An + Xy</i>	36. <i>Hx</i>
7. <i>S-CON</i>	17. <i>COP</i>	27. <i>FAB1</i>	37. <i>An</i>
8. <i>OBS</i>	18. <i>a:p</i>	28. <i>FAB2</i>	38. <i>AG</i>
9. <i>Pairs</i>	19. <i>ES</i>	29. <i>DR1</i>	39. <i>Fd</i>
10. <i>Sum Human</i>	20. <i>Adj. ES</i>	30. <i>DR2</i>	40. <i>Most contents</i>

over two-thirds of the coefficients for the Comprehensive System seem to be missing or unknown. Furthermore, as Aronow, Reznikoff, and Moreland (1995, p. 221) have pointed out, test-retest studies are not a desirable way to assess the Rorschach's reliability. In support of their position, these authors quote the *Standards for Educational and Psychological Testing* (American Psychological Association, 1985, p. 21): "Estimates of stability based on a retest with the same form...may be spuriously inflated due to the effects of memory." Anastasi (1988, p. 118) similarly warns, "For the large majority of psychological tests...retesting with the identical test is not an appropriate technique for finding a reliability."

Aronow and his colleagues argue that the Rorschach may be especially vulnerable to memory effects, so that test-retest studies provide an unrealistic and inflated picture of its true reliability.

### Validity

It is generally acknowledged that the Rorschach has lower overall validity than the Wechsler Intelligence tests (Parker, Hanson, & Hunsley, 1988). The situation is more controversial regarding the MMPI, however. Depending on which meta-analysis is being considered, and how it is interpreted, the mean validity coefficients for the

MMPI are either better than, or roughly equivalent to, those for the Rorschach (Garb, Florio, & Grove, 1998, 1999; Hiller, Rosenthal, Bornstein, Berry, & Brunell-Neulieb, in press; Parker et al., 1988; Parker, Hunsley, & Hanson, 1999). Garb (1999) has provided a thoughtful discussion regarding the limitations of these meta-analyses.

A recent meta-analysis by Hiller and colleagues (in press) has been held out as evidence in support of the Rorschach. As discussed by Garb (1999; see also Garb, Wood, Nezworski et al., 1999), the Hiller meta-analysis is seriously flawed. However, if Rorschach proponents are going to promote it, then all its findings need to be discussed, not just the ones that are favorable to the Rorschach. There were at least four pieces of "bad news for the Rorschach" in Hiller's meta-analysis. First, although the overall validity of the Rorschach and MMPI were not significantly different, it is incorrect to say that Hiller's findings clearly supported the null hypothesis of "no difference." In fact, the overall weighted mean validity coefficients were .37 for the MMPI versus .26 for the Rorschach.

The second piece of "bad news" was that the Rorschach did not correlate highly with self-report instruments. The mean validity coefficient was .28 for the Rorschach. By comparison, the corresponding validity coefficient for the MMPI was .48. Third, and somewhat surprisingly, Hiller and his colleagues (in press) found that the Rorschach had approximately zero correlation with other projective tests. The mean coefficient was .03 for the Rorschach versus .20 for the MMPI. In psychometric terms, the Rorschach showed poor concurrent validity with both self-report and projective tests. Fourth and finally, the Hiller meta-analysis found that the Rorschach did not correlate well with psychiatric diagnoses. The mean validity coefficient was .18 for the Rorschach versus .47 for the MMPI.

This last issue regarding psychiatric diagnoses and the Rorschach seems particularly important, because Rorschach proponents have made some very strong claims. For example, Weiner (1997, pp. 10-11) has argued that the Comprehensive System is useful for the differential diagnosis of several

psychiatric diagnoses, including schizophrenia, depression, and psychopathic personality. Exner (1991, p. 146) has reported that a high score on the Depression Index "correlates very highly with a diagnosis that emphasizes serious affective problems." Levin (1993, pp. 189-190) says that "the Rorschach is ideally suited for assessment of PTSD." Meloy and Gacono (1995, p. 414) claim that the Rorschach is "a sensitive instrument to discriminate between psychopathic and nonpsychopathic subjects."

My colleagues and I recently completed a thorough literature review regarding the relationship of Rorschach scores and diagnoses (Wood, Lilienfeld, Garb, & Nezworski, in press). We asked how many scores would show a well-demonstrated relationship to psychopathology, and by "well demonstrated" we meant that the same effect had been replicated by independent researchers using sound methodology. We reached three main conclusions:

- (1) Patients with schizophrenia, and very possibly patients with Bipolar Disorder and Schizotypal Personality Disorder, tend to give more Deviant Verbalizations, exhibit Bad Form on the Rorschach, and score differently on indexes that incorporate these variables, such as the Thought Disorder Index (Johnston & Holzman, 1979) and the Schizophrenia Index (Exner, 1991, 1993).
- (2) Patients with Borderline PD seem to give an above-average number of Deviant Verbalizations.
- (3) That's all. Otherwise, the Rorschach has not shown a well-demonstrated relationship to these disorders, or to Major Depressive Disorder, Posttraumatic Stress Disorder (PTSD), anxiety disorders other than PTSD, Dissociative Identity Disorder, Dependent, Narcissistic, or Antisocial Personality Disorders, or psychopathy.

### **Internal Consistency and Alternate Forms Reliability**

So far we have reviewed the evidence regarding normative data, interrater reliability, test-retest

reliability, and validity. In none of these areas does the evidence support the claim that the Rorschach is “the single most powerful psychometric instrument ever envisioned.” We will not review evidence regarding the internal consistency of the Rorschach because it is generally accepted that the Rorschach performs poorly in this respect (Stein, 1960, 1962; but see Wagner, Alexander, Roos, & Adair, 1986). Nor can we review evidence of alternate form reliability, because there is no generally accepted alternate form of the test. In conclusion then, we will discuss only two other topics: The factor structure of the Rorschach, and accessibility of the Rorschach Workshops Studies.

### Factor Structure of the Rorschach

As has long been recognized (Cronbach, 1949; Fiske & Baughman, 1953), many Rorschach scores are influenced by the total number of responses that participants give to the test (*R*). The underlying influence of *R* on Rorschach scores is particularly problematic because the number of responses given by a participant is known to be associated with intelligence, education, and socioeconomic status (Anastasi, 1988).

The Comprehensive System for the Rorschach is not exempt from this problem. In a factor analytic study of the Comprehensive System, Meyer (1991, 1992b) concluded that the biggest Rorschach factor was defined by *R*, which accounted for approximately 50% of the common variance in scores (see also Exner, Viglione, & Gillespie, 1984, p. 65). In later research, Meyer (1993) demonstrated that several important Comprehensive System indexes show substantial correlations with response frequency ( $r = .25$  to  $.60$ ).

From a psychometric point of view, the pervasive influence of response frequency on Rorschach scores is highly problematic. As Meyer stated in his first treatment of the topic (1991, but for later apparent modifications see Meyer, 1992a, 1992b, 1997b, 1999), “the great preponderance of variability within the Rorschach data is simply due to the fact that subjects can give as many or as few responses to each card as they like” (p. 225). Meyer (1991, p. 229) concluded:

the traditional use of the Rorschach, where a subject can give as many or as few responses as desired, seriously compromises the validity of the test, as approximately seventy percent of the common variability among Rorschach scores is simply due to error (response frequency). This fact alone calls into question almost all research conducted on the Rorschach, since most studies do not control for this variable.

Besides revealing the problematic influence of *R*, factor analyses have also shown that the intercorrelation pattern of Rorschach scores is often inconsistent with the supposed meaning of the variables. As Meyer (1992b, p. 132) has stated, “the Rorschach’s internal structure does not clearly correspond to that which would be expected from traditional variable interpretation.” He further concludes (p. 133), “it is very doubtful that any theoretical perspective would actually predict the Rorschach structure” that has been found in factor analytic studies of the test. If Meyer’s conclusions are correct, then such studies call into doubt the supposed meaning of many Rorschach scores and raise troubling questions regarding the scores’ construct validity.

### Accessibility of the Rorschach Workshops Studies

As is known to anyone who has read Exner’s books, the unpublished studies of the Rorschach Workshops constitute the main empirical foundation for the Comprehensive System. For example, 157 of Exner’s own works are cited in his books on the Rorschach through 1993. Ninety-nine of these (that is, 63%) are unpublished studies of the Rorschach Workshops. By comparison, 27 are journal articles. Thus the ratio of Workshops Studies to journal articles is more than three to one.

In 1993 and 1994, one of us (J.M.W.) and three of his colleagues tried to obtain reprints or preprints of the Workshops Studies to review their methodology and results. Four of us, all PhD level psychologists, each wrote a letter to Exner. Altogether we requested copies of 23 studies, and offered to pay for copying and mailing costs. In response, we

received letters from the Rorschach Workshops. We were told that some studies were not in their files, others had not been written in publishable form. No reprints or preprints for Workshops Studies were sent to us, nor were their Methods or Results sent to us, nor were there any offers to send them to us. Instead we were told "the works that you've requested are not available in a form that can be easily copied and forwarded to you." We were told that raw data might be available, but that we might have to pay for computer costs. However, we saw no point in paying for data if we could not obtain the studies' Methods and Results.

It might have been unreasonable if we'd requested Exner's data after 20 years. But that's not what we did. His recent books contain numerous citations to the unpublished Workshops studies. So we requested reprints or preprints of these studies, not the data. If copies of the Workshops studies cannot be released to other scholars, we question whether they should be cited to support the reliability and validity of the Comprehensive System. Moreover, the inaccessibility of such studies renders it difficult to exclude the possibility that certain reported correlates of Rorschach indexes are attributable to file-drawer effects, i.e., the tendency for negative results to remain unpublished (Rosenthal, 1991).

In a response to our criticisms, Exner (1996) has stated that some of the Workshops studies have been rewritten and published as peer-reviewed articles. This is true, but apparently only in a limited number of cases. For example, we count 76 separate unpublished Workshops studies that are cited in Exner books published in 1991 and 1993. Few or none of these studies seem to have been rewritten and published as articles. Exner's books continue to cite a large number of Workshops Studies that have never been published in journals or peer-reviewed.

### Conclusions

We have reviewed the psychometric qualities of the Rorschach one by one. The following five

conclusions seem in order and are consistent with other recent critiques of the test (Dawes, 1994; Gann, 1995; Garb, 1998, 1999; Garb, Wood, & Nezworski, 1999; Garb, Wood, Nezworski et al., 1999; Hunsley & Bailey, in press; Sechrest, Stickle, & Stewart, 1998; Wood et al., in press; Wood et al., 1996a, 1996b; Wood, Nezworski, Stejskal, Garven, & West, 1999; but for more positive views of the Rorschach see Exner, 1996; Ganellen, 1996a, 1996b; Ornberg & Zalewski, 1994; Stricker & Gold, in press; Viglione, in press; Weiner, 1996, 1997).

First, the norms for the Comprehensive System represent a step forward for the Rorschach. However, these norms are older and based on substantially smaller samples than the norms for the WAIS-III and MMPI-2. The lack of culturally sensitive norms for the Rorschach has been repeatedly criticized. Furthermore, current Rorschach norms for children appear to be outdated and based on unrepresentative samples.

Second, some Comprehensive System scores have good or even excellent interrater reliability. However, median reliability appears to be under .85 and well below the reliability of the WAIS-III subtests. The claim that all Comprehensive System scores have a reliability above .85 seems to be a clear overstatement. Third, test-retest reliability is apparently unknown for about two-thirds of Comprehensive System variables. Claims that the Comprehensive System has "excellent" test-retest reliability are premature. Fourth, claims that the Comprehensive System is a useful clinical measure of depression, PTSD, or psychopathy also appear to be overstatements. In fact, very few Rorschach scores bear a well-demonstrated relationship to psychiatric diagnoses.

Fifth and finally, taken altogether, the psychometric evidence does not support the claim that the Rorschach is "the single most powerful psychometric instrument ever envisioned." In fact, this claim, published in psychology's premier journal by the APA Board of Professional Affairs (1998), may be as much an overstatement as the old assertion that the Rorschach is an x-ray of the mind.



## References

- Acklin, M. W., McDowell, C. J., & Verschell, M. S. (in press). Interobserver agreement, intraobserver reliability, and the Rorschach Comprehensive System. *Journal of Personality Assessment*.
- American Psychological Association. (1985). *Standards for educational and psychological testing*. Washington, DC: Author.
- American Psychological Association. (1993). Guidelines for providers of psychological services to ethnic, linguistic, and culturally diverse populations. *American Psychologist*, 48, 45-48.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: MacMillan.
- Aposhian, M. A. (1995). Iranians and the Rorschach: A question of cross-cultural validity (Doctoral dissertation, The Wright Institute, 1994). *Dissertation Abstracts International*, 55, 3633B.
- Aronow, E., Reznikoff, M., & Moreland, K. L. (1995). The Rorschach: Projective technique or psychometric test? *Journal of Personality Assessment*, 64, 213-228.
- Baca, B. J. (1994). A comparison of selected Rorschach variables of Hispanic pedophiles and Hispanic normals with each other, and with nonpatient normals and character disorders: Some issues of acculturation (Doctoral dissertation, The Union Institute, 1993). *Dissertation Abstracts International*, 54, 5934B.
- Board of Professional Affairs. (1998). Awards for distinguished professional contributions: John Exner. *American Psychologist*, 53, 391-392.
- Butcher, J. N., Nezami, E., & Exner, J. E. (1998). Psychological assessment of people in diverse cultures. In S. S. Kazarian & D. R. Evans (Eds.), *Cultural clinical psychology* (pp. 61-105). New York: Oxford University Press.
- Constantino, G., Flanagan, R., & Malgady, R. (1995). The history of the Rorschach: Overcoming bias in multicultural projective assessment. *Rorschachiana*, 20, 148-171.
- Cronbach, L. J. (1949). Statistical methods applied to Rorschach scores: A review. *Psychological Bulletin*, 46, 393-429.
- Cuellar, I. (1998). Cross-cultural clinical psychological assessment of Hispanic Americans. *Journal of Personality Assessment*, 70, 71-86.
- Dana, R. H. (1993). *Multicultural assessment perspectives for professional psychology*. Boston: Allyn & Bacon.
- Dana, R. H. (1998). Cultural identity assessment of culturally diverse groups: 1997. *Journal of Personality Assessment*, 70, 1-16.
- Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on myth*. New York: Free Press.
- Day, R., Boyer, L. B., & Devos, G. A. (1989). Progressive constriction in Apache youth. In G. A. Devos & L. B. Boyer (Eds.), *Symbolic analysis cross-culturally: The Rorschach Test* (pp. 293-334). Berkeley: University of California Press.
- Devos, G. A. (1989). Personality continuities and cultural change in Japanese Americans. In G. A. Devos & L. B. Boyer (Eds.), *Symbolic analysis cross-culturally: The Rorschach Test* (pp. 93-136). Berkeley: University of California Press.
- Devos, G. A., & Miner, H. (1989). Oasis and Casbah: Acculturative stress. In G. A. Devos & L. B. Boyer (Eds.), *Symbolic analysis cross-culturally: The Rorschach Test* (pp. 201-245). Berkeley: University of California Press.
- 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. (1980). But it's only an inkblot. *Journal of Personality Assessment*, 44, 563-577.
- Exner, J. E. (1986). *The Rorschach: A Comprehensive System. Volume 1: Basic foundations* (2nd ed.). New York: Wiley.
- Exner, J. E. (1988). Problems with brief Rorschach protocols. *Journal of Personality Assessment*, 52, 640-647.
- 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. (1996). A comment on "The Comprehensive System for the Rorschach: A critical examination." *Psychological Science*, 7, 11-13.
- Exner, J. E., Armbruster, G. L., & Viglione, D. (1978). The temporal stability of some Rorschach features. *Journal of Personality Assessment*, 42, 474-482.
- Exner, J. E., Thomas, E. A., & Mason, B. (1985). Children's Rorschachs: Description and prediction. *Journal of Personality Assessment*, 49, 13-20.
- Exner, J. E., Viglione, D. J., & Gillespie, R. (1984). Relationships between Rorschach variables as relevant to the interpretation of structural data. *Journal of Personality Assessment*, 48, 65-70.
- 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.
- Fiske, D. W., & Baughman, E. E. (1953). Relationships between Rorschach scoring categories and the total number of responses. *Journal of Abnormal and Social Psychology*, 48, 25-32.
- Frank, G. (1992). The response of African Americans to the Rorschach: A review of the literature. *Journal of Personality Assessment*, 59, 317-325.
- Frank, G. (1993). The use of the Rorschach with Hispanic Americans. *Psychological Reports*, 72, 276-278.
- Gann, M. K. (1995). The Rorschach and other projective methods. In J. Ziskin (Ed.), *Coping with psychiatric and psychological testimony. Vol. II: Special topics* (5th ed., pp. 823-884). Los Angeles: Law and Psychology Press.
- 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: Erlbaum.

- Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment*. Washington, DC: American Psychological Association.
- Garb, H. H. (1999). Call for a moratorium on the use of the Rorschach Inkblot Test. *Assessment, 6*, 3xx-3xx.
- Garb, H. N., Florio, C. M., & Grove, W. M. (1998). The validity of the Rorschach and the Minnesota Multiphasic Personality Inventory: Results from meta-analyses. *Psychological Science, 9*, 402-404.
- Garb, H. N., Florio, C. M., & Grove, W. M. (1999). The Rorschach controversy: Reply to Parker, Hunsley, and Hanson. *Psychological Science, 10*, 293-294.
- Garb, H. N., Wood, J. M., & Nezworski, M. T. (1999). *Projective techniques and the detection of child sexual abuse*. Manuscript submitted for publication.
- Garb, H. N., Wood, J. M., Nezworski, M. T., Grove, W. M., & Stejskal, W. J. (1999). *Towards a resolution of the Rorschach controversy*. Manuscript submitted for publication.
- Glass, M. H., Bieber, S. L., & Tkachuk, M. J. (1996). Personality styles and dynamics of Alaska native and non-native incarcerated men. *Journal of Personality Assessment, 66*, 583-603.
- Gray-Little, B. (1995). The assessment of psychopathology in racial and ethnic minorities. In J. N. Butcher (Ed.), *Clinical personality assessment: Practical approaches* (pp. 140-157). New York: Oxford University Press.
- Gray-Little, B., & Kaplan, D. A. (1998). Interpretation of psychological tests in clinical and forensic evaluations. In J. Sandoval, C. L. Frisby, K. F. Geisinger, J. D. Scheuneman, & J. R. Grenier (Eds.), *Test interpretation and diversity* (pp. 141-178). Washington, DC: American Psychological Association.
- Groth-Marnat, G. (1990). *Handbook of psychological assessment* (2nd ed.). New York: John Wiley & Sons.
- Haller, N., & Exner, J. E. (1985). The reliability of Rorschach variables for inpatients presenting symptoms of depression and/or helplessness. *Journal of Personality Assessment, 49*, 516-521.
- Hiller, J. B., Rosenthal, R., Bornstein, R. F., Berry, D. T. R., & Brunell-Neuleib, S. (in press). A comparative meta-analysis of Rorschach and MMPI validity. *Psychological Assessment*.
- Howes, R. D., & DeBlassie, R. R. (1989). Modal errors in the cross cultural use of the Rorschach. *Journal of Multicultural Counseling and Development, 17*, 68-84.
- Hunsley, J., & Bailey, J. M. (in press). The clinical utility of the Rorschach: Unfulfilled promises and an uncertain future. *Psychological Assessment*.
- Johnston, M., & Holzman, P. S. (1979). *Assessing schizophrenic thinking*. San Francisco: Jossey-Bass.
- Krall, V., Sachs, H., Lazar, B., Rayson, B., Grove, G., Novar, L., & O'Connell, L. (1983). Rorschach norms for inner city children. *Journal of Personality Assessment, 47*, 155-157.
- Levin, P. (1993). Assessing posttraumatic stress disorder with the Rorschach projective technique. In J. P. Wilson & B. Raphael (Eds.), *International handbook of traumatic stress syndromes* (pp. 189-200). New York: Plenum Press.
- McDowell, C., & Acklin, M. W. (1996). Standardizing procedures for calculating Rorschach interrater reliability: Conceptual and empirical foundations. *Journal of Personality Assessment, 66*, 308-320.
- Meloy, J. R., & Gacano, C. B. (1995). Assessing the psychopathic personality. In J. N. Butcher (Ed.), *Clinical personality assessment; Practical approaches* (pp. 410-422). New York: Oxford University Press.
- Meyer, G. J. (1991). An empirical search for fundamental personality and mood dimensions within the Rorschach test (Doctoral dissertation, Loyola University of Chicago, 1991). *Dissertation Abstracts International, 52*, 1071B-1072B.
- Meyer, G. J. (1992a). Response frequency problems in the Rorschach: Clinical and research implications with suggestions for the future. *Journal of Personality Assessment, 58*, 231-244.
- Meyer, G. J. (1992b). The Rorschach's factor structure: A contemporary investigation and historical review. *Journal of Personality Assessment, 59*, 117-136.
- Meyer, G. J. (1993). The impact of response frequency on the Rorschach constellation indices and on their validity with diagnostic and MMPI-2 criteria. *Journal of Personality Assessment, 60*, 153-180.
- Meyer, G. J. (1997a). Assessing reliability: Critical corrections for a critical examination of the Rorschach Comprehensive System. *Psychological Assessment, 9*, 480-489.
- Meyer, G. J. (1997b). On the integration of personality assessment methods: The Rorschach and MMPI. *Journal of Personality Assessment, 68*, 297-330.
- Meyer, G. J. (1997c). Thinking clearly about reliability: More critical corrections regarding the Rorschach Comprehensive System. *Psychological Assessment, 9*, 495-498.
- Meyer, G. J. (1999). The convergent validity of MMPI and Rorschach Scales: An extension using profile scores to define response and character styles on both methods and a reexamination of simple Rorschach response frequency. *Journal of Personality Assessment, 72*, 1-35.
- Moon, T. I., & Cundick, B. P. (1983). Shifts and constancies in Rorschach responses as a function of culture and languages. *Journal of Personality Assessment, 47*, 345-349.
- Munson, R. J. (1993). A cross-cultural study of selected Rorschach indices of Chinese and American normal and schizophrenic subjects (Doctoral dissertation, The University of Tennessee, 1992). *Dissertation Abstracts International, 54*, 2216B.
- Okazaki, S., & Sue, S. (1995). Cultural considerations in psychological assessment of Asian-Americans. In J. N. Butcher (Ed.), *Clinical personality assessment: Practical approaches* (pp. 107-119). New York: Oxford University Press.
- Ornberg, B., & Zalewski, C. (1994). Assessment of adolescents with the Rorschach: A critical review. *Assessment, 1*, 209-217.
- Parker, K. C. H., Hanson, R. K., & Hunsley, J. (1988). MMPI, Rorschach, and WAIS: A meta-analytic comparison of reliability, stability, and validity. *Psychological Bulletin, 103*, 367-373.
- Parker, K. C. H., Hunsley, J., & Hanson, R. K. (1999). Old wine from old skins sometimes tastes like vinegar: A response to Garb, Florio, and Grove. *Psychological Science, 10*, 291-292.

- Psychological Corporation. (1997). *Wechsler Adult Intelligence Scale, Third Edition. Wechsler Memory Scale, Third Edition. Technical Manual*. San Antonio, TX: Author.
- Ritzler, B. A. (1996). Projective methods for multicultural personality assessment. In L. A. Suzuki, P. J. Meller, & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment* (pp. 115-135). San Francisco: Jossey-Bass.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.
- Sanchez, E. M. (1993). Affective scoring of Native American female Rorschach protocols (Doctoral dissertation, Miami Institute of Psychology of the Caribbean Center for Advanced Studies, 1992). *Dissertation Abstracts International*, 53, 6027B.
- Sangro, F. M. (1997). Location tables, form quality, and popular responses in a Spanish sample of 470 subjects. *Rorschachiana*, 22, 38-66.
- Schmidt, F. L. (1988). The problem of group differences in ability test scores in employment selection. *Journal of Vocational Behavior*, 32, 272-292.
- 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.
- Stein, H. (1960). Rotation and reliability of the Rorschach. *Journal of Projective Techniques*, 24, 171-181.
- Stein, H. (1962). An analysis of two components entering into Rorschach reliability values. *Journal of Projective Techniques*, 26, 474-477.
- Stricker, G., & Gold, J. R. (in press). The Rorschach: Towards a nomothetically based, idiographically applicable, configurational model. *Psychological Assessment*.
- Velasquez, R. J. (1995). Personality assessment of Hispanic clients. In J. N. Butcher (Ed.), *Clinical personality assessment: Practical approaches* (pp. 120-139). New York: Oxford University Press.
- Velasquez, R. J., & Callahan, W. J. (1992). Psychological testing of Hispanic Americans in clinical settings: Overview and issues. In K. F. Geisinger (Ed.), *Psychological testing of Hispanics* (pp. 253-265). Washington, DC: American Psychological Association.
- Viglione, D. J. (in press). A review of recent research addressing the utility of the Rorschach. *Psychological Assessment*.
- Wagner, E. E., Alexander, R. A., Roos, G., & Adair, H. (1986). Optimum split-half reliabilities for the Rorschach: Projective techniques are more reliable than we think. *Journal of Personality Assessment*, 50, 107-112.
- 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.
- Wood, J. M., Lilienfeld, S. O., Garb, H. N., & Nezworski, M. T. (in press). The Rorschach test in clinical diagnosis: A critical review, with a backward look at Garfield (1947). *Journal of Clinical Psychology*.
- 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. (1997). The reliability of the Comprehensive System for the Rorschach: A comment on Meyer (1997). *Psychological Assessment*, 9, 490-494.
- Wood, J. M., Nezworski, M. T., Stejskal, W. J., Garven, S., & West, S. G. (1999). Methodological issues in evaluating Rorschach validity: A comment on Burns and Viglione (1996), Weiner (1996), and Ganellen (1996). *Assessment*, 6, 115-120.