

WPE WebPsychEmpiricist

The Rorschach's false positive rate is 81%

R. K. McKinzey, Ph.D.

5/15/05

Oakland, CA

Abstract

Two different studies sent normal Rorschach protocols to qualified psychologists. They found that over 80% of the psychologists misidentified the normal protocols as showing severe psychopathology. Using the Comprehensive System did not improve the false positive rate, which stands as 81%. The two studies are described in detail, and their result tables reproduced. This review is necessary since one study is a dissertation not readily available, and my citation of it elsewhere has been challenged.

The Rorschach's false positive rate is 81%

All tests are bound by the rigid rules of the 2x2 table (Table 1). To be useful, the test must distinguish between the condition it was designed to measure and all other conditions. The test's ability to measure the former is called its sensitivity, and its ability to measure the latter is called its specificity. When a test fails to identify a condition it is designed to measure, the failure is called a false negative. When it mistakenly identifies a condition that is not present, the mistake is called a false positive. The percentage of false negatives and false positives in a given population are called the false negative rate (FNR) and false positive rate (FPR). Tests of psychological abnormality must distinguish between people with an abnormality and those who are normal. A psychological test that calls normal people abnormal is said to have a high false positive rate.

Elsewhere (Brodsky & McKinzey, 2002; McKinzey & Campagna, 4/27/02) I have stated that the Rorschach's false positive rate is 81%, citing a dissertation (Mittman, 1983). A review paper (Hilsenroth & Stricker, 2004) has challenged that citation, calling it unjustified, apparently without having read the dissertation. (Since the dissertation was never produced as a journal article, a common outcome for dissertations, the citation has to be purchased.¹) Therefore, I will describe the basis for my assertion.

The Rorschach was once thought to be impervious to being faked (Albert, Fox, & Kahn, 1980). Albert et al. therefore collected a group of Rorschach protocols from six people with psychoses, six college students taking the test under standard instructions, six uninformed college students asked to fake the Rorschach as if they were psychotic, and six college students informed about the signs and symptoms of psychosis. All the college students had normal MMPIs and "were not seriously disturbed psychologically." The authors then got 46 psychologists to review an unscored protocol from each subgroup, and offer a categorization of each protocol. All the psychologists were Fellows of the Society for Personality Assessment, "considered as a group to probably represent the highest existing level of Rorschach proficiency." (p. 116) Table 2 presents their results.

Their first finding was that the Rorschach could be faked, and such faking is rarely detected. Then, "A secondary but rather disturbing finding is that the expert judges rated 24% of the Normal protocols as psychotic, while correctly identifying psychosis in only 48% of the actual Psychotic protocols...The findings of this study as a whole seem to call into serious question the validity of the Rorschach procedure for diagnostic purposes." (p. 118) More specifically, 46 qualified

¹ Dissertations can be purchased for \$38 from UMI's Dissertation Express: <http://www.umi.com/umi/dissertations/disexpress.shtml>. Mittman's dissertation's number is 8315540. Links to UMI can be found at this article's WPE abstract page. The question of whether this research was suppressed by the Rorschach industry is not yet resolved.

psychologists reviewed one of 4 normal protocols. Eight of the resulting judgments were correct. In other words, the specificity was 17%, and the FPR was 83%.

However, Exner's Comprehensive System (CS) was new, and only one system among many in use by psychologists.² In doing her doctoral dissertation³, Beatrice Landman Mittman decided to replicate Albert et al. (1980) using only psychologists who used the CS and to send them scored protocols.

In doing so, Mittman obtained the help of her coauthor (Exner, Armbruster, & Mittman, 1978) and employer John Exner. Together, and on Rorschach Workshops letterhead⁴, they recruited psychologists who a) were in "teaching, supervisory, or senior positions" and b) alumnae of the Rorschach Workshops that taught them to use the CS. The Workshops were taught by either Exner or the coauthor of the CS, Irving Weiner. Mittman got 90 such psychologist participants.

Mittman recruited 18 adult office workers to take the Rorschach. They were nonpatients who had normal MMPIs, and thus considered Normal. Nine trained graduate students administered the test and scored it. Six took the test under standard conditions. Six were asked to fake psychosis on the test, but not informed about the nature of psychosis. Six were asked to fake psychosis, and were informed about the condition. Six protocols of people with schizophrenia and six of people with depression were drawn from the CS standardization pool.

Each of the 90 psychologists were then sent five protocols and asked to categorize them. The packet of protocols contained a randomly varied mixture of the five groups. The 90 psychologists thus made 5 judgments each, for a total of 450 judgments.

Table 3 (taken from Mittman's Table 4, p.45) presents the results. As with Albert et al (1980), the Rorschach was easily malingered and rarely detected⁵. 57% of the schizophrenic protocols, 47% of the depressed protocols, and 19% of the normal protocols were correctly identified. Of the 170 protocols with either of the two psychiatric conditions, 89 were identified, making the sensitivity 52%. The specificity is 19% AND THE FALSE POSITIVE RATE IS 81%!

In the discussion section, Mittman notes the high FPR: "Many of the judges saw the control protocols as displaying a higher degree of psychopathology than would be expected in the normal population...Perhaps the Rorschach is not as effective a tool for diagnostic labeling as clinicians tend to believe." (p.63) On the other hand, "Judges may have been able to pick up indications of pathology from the Rorschach which were not detectable from the self-report MMPI." (p. 64) Maybe "the bias toward psychopathology was due more to the expectations of the judges than to a structural bias in the Rorschach. (p.65) After all, "Few human beings are

² See (Wood, Nezworski, Lilienfeld, & Garb, 2003) for a review of the history of the Rorschach's several interpretative systems.

³ She was getting her doctorate from Long Island University.

⁴ Mittman's Appendix A is a copy of the letter. She is described as the "Assistant to the Director", Exner.

totally devoid of symptoms, and judges may have preferred to note these tendencies rather than to ignore them entirely.” (p. 65) However, “It is unclear whether judges failed to label control protocols as normal due to the inconsistent definition of normal and of the other diagnostic labels available to them or due to the Rorschach’s bias toward psychopathology.” (p. 66)

Arguing that the Rorschach identified unsuspected pathology in this carefully screened sample ignores the seriousness of the pathology identified. The judges thought 18% of this sample had a Borderline or Schizoid Personality Disorder, and that 25% were seriously depressed. The judges thought 2% were faking, and another 2% had a drug-induced reaction. Such an argument strains the credulity, especially in light of reviews showing the Rorschach CS wildly overpathologizes in a variety of ways (Wood, Nezworski, Garb, & Lillienfeld, 2001; Wood et al., 2003). The fact remains: The Rorschach’s false positive rate is 81% (Albert et al., 1980; Mittman, 1983).

⁵ This point was not conceded by Exner until 1991 (Exner & Weiner, 1978, 1991).

References

- Albert, S., Fox, H. M., & Kahn, M. W. (1980). Faking psychosis on the Rorschach: Can expert judges detect malingering? *Journal of Personality Assessment*, *44*(2), 115-119.
- Brodsky, S. L., & McKinzey, R. K. (2002). The Ethical Confrontation of the Unethical Forensic Colleague. *Professional Psychology: Research & Practice*, *33*(3), 307-309.
- Exner, J. E., Armbruster, G., & Mittman, B. (1978). The Rorschach response process. *Journal of Personality Assessment*, *42*(1), 27.
- Exner, J. E., & Weiner, I. B. (1978, 1991). *The Rorschach: A comprehensive system (Vol. II)*. New York: Wiley.
- Hilsenroth, M. J., & Stricker, G. (2004). A consideration of challenges to psychological assessment instruments used in forensic settings: Rorschach as exemplar. *Journal of Personality Assessment*, *83*(2), 141.
- McKinzey, R. K., & Campagna, V. (4/27/02). The Rorschach, Exner's Comprehensive System, Interscorer Agreement, and Death. *WebPsychEmpiricist*. Retrieved April 27, 2002, from http://www.wpe.info/papers_table.html
- Mittman, B. L. (1983). Judges' ability to diagnose schizophrenia on the Rorschach: The effect of malingering. *Dissertation Abstracts International*, 44-04B, Page 1248. (UMI No. 8315540)
- Wood, J. M., Nezworski, M. T., Garb, H. N., & Lillienfeld, S. O. (2001). The misperception of psychopathology: Problems with the norms of the Comprehensive System for the Rorschach. *Clinical Psychology: Science and Practice*, *8*(3), 350-373.
- Wood, J. M., Nezworski, M. T., Lilienfeld, S. O., & Garb, H. N. (2003). *What's Wrong With the Rorschach? Science Confronts the Controversial Inkblot Test*. New York: Wiley.

Table 1

Test accuracy: The 2x2 Table

		Criterion		
		Normal	Abnormal	
Test results	Normal	True Negative/ Specificity	False Negative	Negative Predictive Power (NPP)
	Abnormal	False positive	True Positive/ Sensitivity	Positive Predictive Power (PPP)
		Total Normals	Total Abnormals	Total in Sample

Specificity = True Negative/ Total Normals: Identification of those without the condition

Sensitivity = True positives/Total Abnormals: Identification of those with the condition

False Negative = 1-Sensitivity: How many with the condition does the test misidentify?

False Positive = 1-Specificity : How many without the condition does misidentify?

Hit rate = True Negatives + True Positives/Total in Sample: Overall accuracy

Base rate = Total Abnormals/Total in Sample: How often does the condition occur?

Positive Predictive Power (also Positive Predictive Value) = True test Positives/Total test

Positives: Given the Base Rate, how useful is the test in identifying true positives?

Negative Predictive Power (also Negative Predictive Value) = True test Negatives/Total test

Negatives: Given the Base Rate, how useful is the test in identifying true negatives?

Table 2

Frequencies of General Diagnoses Given to Each Type of Protocol

Diagnosis given	Type of protocol			
	Psychotic	Informed Faker	Uninformed Faker	Normal
Psychosis	22 (48%)	33 (72%)	21 (46%)	11 (24%)
Organic Brain Syndrome	2 (4%)	0 (0%)	2 (4%)	1 (2%)
Character Disorder	8 (17%)	6 (13%)	7 (15%)	6 (15%)
Neurosis	8 (17%)	2 (4%)	6 (13%)	17 (37%)
Transient Situational Disturbance	0 (0%)	0 (0%)	3 (7%)	2 (4%)
Essentially normal	3 (7%)	1 (2%)	4 (9%)	8 (17%)
Faked	2 (4%)	4 (9%)	3 (7%)	1 (2%)
Other	1 (2%)	0 (0%)	0 (0%)	0 (0%)
Total	46	46	46	46

Note: Taken from Albert et al. (1980), Table 1, p.117. Percentages are rounded.

Table 3

Frequency of Diagnoses Made by Psychologists for Each Type of Protocol

Diagnosis	Type of Protocol					Total
	Schizophrenic	Uninformed fake	Informed fake	Depressed	Normal	
Judges' diagnosis						
Schizophrenic	51 (57%)	5 (5.5%)	16 (18%)	12 (15%)	0 (0%)	84
Endogenous Depression	2 (2%)	13 (14%)	3 (3%)	16 (20%)	11 (11%)	45
Reactive Depression	1 (1%)	8 (9%)	9 (10%)	10 (12%)	23 (23%)	51
Manic Depression	9 (10%)	2 (2%)	3 (3%)	12(15%)	1 (1%)	27
Fake	1 (1%)	1 (1%)	1 (1%)	1 (1%)	2 (2%)	6
Normal	0 (0%)	18 (20%)	4 (4%)	0 (0%)	19(19.2%)	41
Obsessive Personality	0 (0%)	10 (11%)	5 (6%)	1 (1)	4 (4%)	20
Borderline Personality	12 (13.5%)	6 (7%)	17 (19%)	14(17%)	5 (5%)	54
Inadequate Personality	1 (1%)	14 (15%)	4 (4%)	6 (6%)	13 (13%)	37
Schizoid Personality	6 (7%)	4 (4%)	6 (7%)	1 (1%)	13 (13%)	30
Hysterical Personality	3 (3%)	7 (8%)	18 (20%)	3 (4%)	6 (6%)	37
Drug Induced Reaction	3 (3%)	3 (3%)	4 (4%)	6 (7%)	2 (2%)	18
Total	89	91	90	81	99	450

Note. Taken from Mittman (1983), Table 4, p. 45. Percentages are rounded.