THE VALIDITY OF THE RORSCHACH AND THE MINNESOTA MULTIPHASIC PERSONALITY INVENTORY: Results From Meta-Analyses

Howard N. Garb, Colleen M. Florio, and William M. Grove
1Pittsburgh VA Health Care System; 2Department of Psychiatry, University of Pittsburgh; 3Gynecologic Oncology Group, Roswell Park Cancer Institute, Buffalo, New York; and 4Department of Psychology, University of Minnesota

Abstract—Results from meta-analyses have been widely cited to defend the validity of the Rorschach. However, the meta-analyses have been flawed. For example, one meta-analysis included results that were obtained by calculating correlations but not results that were obtained by conducting t tests or analyses of variance. When we reanalyzed the data from the most widely cited meta-analysis (Parker, Hanson, & Hunsley, 1988), we found that for confirmatory studies (also called convergent-validity studies), the Minnesota Multiphasic Personality Inventory [MMPI] explained 23% to 30% of the variance, whereas the Rorschach explained only 8% to 13% of the variance. These results indicate that the Rorschach is not as valid as the MMPI.

The use of the Rorschach Inkblot Method continues to be controversial (e.g., Exner, 1996; Wood, Nezworski, & Stejskal, 1996a, 1996b). However, results from several meta-analyses seem to support the Rorschach (Atkinson, 1986; Atkinson, Quarrington, Alp, & Cyr, 1986; Parker, 1983; Parker, Hanson, & Hunsley, 1988). For example, based on the results for one of the meta-analyses, Parker et al. (1988) concluded that the convergent-validity estimates for the Rorschach and MMPI were not significantly different. . . . It appears that both the MMPI and Rorschach can be considered to have adequate psychometric properties if used for the purposes for which they were designed and validated. (abstract, p. 367)

The results from the meta-analyses, especially the results from the meta-analysis conducted by Parker et al. (1988), have been widely cited to defend the validity of the Rorschach (e.g., Beutler & Davison, 1995; Ganellen, 1996; Masling, 1997; Shontz & Green, 1992; Weiner, 1996, 1997). It does seem unfair to argue that the Rorschach should not be used in clinical and forensic practice if meta-analyses have supported its use, especially in view of the fact that the meta-analysis conducted by Parker et al. (1988) was published in a highly prestigious journal (Psychological Bulletin) and has been reprinted as an illustrative example of meta-analysis in a book on methodology (Kazdin, 1992).

Unfortunately, the meta-analyses that have been cited to support the Rorschach have been flawed. For example, in one meta-analysis, results on validity were not analyzed separately from results on reliability (Parker, 1983); in another, meta-analysis estimates were not made of the magnitude of effects (Atkinson et al., 1986); and in a third, meta-analysis effect sizes were estimated but statistical proce-

Address correspondence to Howard N. Garb, Behavioral Health (116A-H), VA Medical Center, 7180 Highland Dr., Pittsburgh, PA 15206-1297; e-mail: garb.howard@pittsburgh.va.gov.
effect size meant that a test was valid. For example, in one study (Hersen & Greaves, 1971), a small effect size indicated that Rorschach productivity is not related to verbal reinforcement. Similarly, in many studies (e.g., Griffin, Finch, Edwards, & Kendall, 1976; Newmark, Newmark, & Cook, 1975), an abbreviated form of the MMPI was highly correlated with the standard form of the MMPI, but the effect size was estimated to be zero (apparently because statistical tests between the short form and regular form were typically nonsignificant).

**REANALYSIS OF THE PARKER ET AL. (1988) DATA**

**Method**

The terms “unknown validity” and “convergent validity” are misleading. For example, it seems odd to report a validity coefficient for an unknown-validity category (if validity really were unknown, one would not be able to report a validity coefficient). Instead of referring to unknown-validity studies and convergent-validity studies, we refer here to exploratory studies and confirmatory studies. We coded a study as being exploratory if Parker et al. had coded it as being of unknown validity, and we coded a study as being confirmatory if Parker et al. had coded it as measuring convergent validity.

In a meta-analysis, an effect-size estimate is obtained for each study, and statistical procedures are used to aggregate the estimates across studies. We used the effect-size estimates that Parker et al. had calculated. They calculated $r^2$ values for studies that used correlation- or chi-square analyses and $\omega^2$ values for studies that used $t$ tests and ANOVAs.

Multiple findings from a single study may be nonindependent. To control for nonindependence, we included only a single estimate of effect size from each study. As did Parker et al., we used the median value in our meta-analysis when more than one estimate of effect size could be calculated. If there was an even number of effect sizes that could be calculated, we took the average between the two middle values (weighted by the sample sizes for the two values). Parker et al. had randomly selected the higher or lower value at the median.

Having obtained our median estimates of effect size, we used the same statistical procedure that Parker et al. used to aggregate the results. After taking the square root of every median estimate of effect size, we transformed all of these values to be unbiased estimators of the population correlation coefficient by using Equation 3 from Hedges and Olkin’s (1985, p. 225) book. All of these values were then transformed into Fisher $z$ scores to normalize the distribution. Weighted mean Fisher $z$ scores were calculated, and these scores were transformed into correlation coefficient equivalents.

**Results**

**Confirmatory studies**

The results for the confirmatory data indicate that the MMPI is more valid than the Rorschach. Correlation coefficient equivalents of the average $z$ scores were .29 for the Rorschach and .48 for the MMPI. The difference between the two correlation coefficients was statistically significant, $z = 8.04, p < .001$. Corresponding $r^2$ values were .08 and .23 for the Rorschach and the MMPI, respectively. Although Parker et al. based their confirmatory- (convergent-) validity results on only 5 studies for the Rorschach and 31 studies for the MMPI, we based our results on 18 studies for the Rorschach and 66 studies for the MMPI.

In a second meta-analysis, studies were excluded if a Rorschach score was used as a construct measure in a Rorschach study, a self-report personality measure was used in an MMPI study, or a small effect indicated that validity was good. Using these exclusion rules, we eliminated the results from 4 Rorschach studies and 30 MMPI studies. Thus, the results were based on 14 Rorschach studies and 36 MMPI studies. These results also indicate that the MMPI is more valid than the Rorschach. Correlation coefficient equivalents of the average $z$ scores were .33 for the Rorschach and .55 for the MMPI. The difference between the two correlation coefficients was statistically significant, $z = 8.25, p < .001$. Corresponding $r^2$ values were .11 and .30 for the Rorschach and the MMPI.

We conducted a third meta-analysis after eliminating studies that used chi-square analyses. Chi-square analyses were used in four of the Rorschach confirmatory studies, but they were not used in any of the MMPI studies. When chi-square analyses were conducted, the dependent variables were continuous, but they were turned into dichotomous variables by the investigators. They may have been turned into dichotomous variables because by looking at the data while they set a cutoff score, the researchers might capitalize on chance. When the effect-size estimates based on the chi-square analyses were excluded, the correlation coefficient equivalents of the average $z$ scores were .36 for the Rorschach and .55 for the MMPI. The difference between the two correlation coefficients was statistically significant, $z = 5.52, p < .001$. Corresponding $r^2$ values were .13 and .30 for the Rorschach and the MMPI.

**Exploratory studies**

Though Parker et al. emphasized the importance of the results from the confirmatory studies, for the sake of completeness, we also reanalyzed the results from the exploratory studies. For 24 studies on the Rorschach and 138 studies on the MMPI, correlation coefficient equivalents of the average $z$ scores were .11 for the Rorschach and .11 for the MMPI. The difference between the two correlation coefficients was not statistically significant, $z = 0, p > .05$. Corresponding $r^2$ values were .01 and .01 for the Rorschach and the MMPI, respectively.

As we did for the confirmatory studies, we conducted a second meta-analysis to (a) control for shared methodological variance between the assessment instruments and the criteria and (b) ensure that effect size was not negatively related to validity. These exclusion rules eliminated the results from 9 Rorschach studies and 63 MMPI studies. Thus, the results were based on 15 Rorschach studies and 75 MMPI studies. Correlation coefficient equivalents of the average $z$ scores were .18 for the Rorschach and .11 for the MMPI. The difference between the two correlation coefficients was statistically significant, $z = -2.08, p < .05$. Corresponding $r^2$ values were .03 and .01 for the Rorschach and the MMPI, respectively.

**Rorschach scoring systems**

We also compared different Rorschach scoring systems. Studies were included only if (a) the criterion measure was not a Rorschach score and (b) validity was positively related to effect size. Exner’s Comprehensive System (Exner, 1974, 1978) was used in 2 of the 14 confirmatory-validity studies and 3 of the 15 exploratory studies, and Klopfers scoring system (Klopfers, Ainsworth, Klopfers, & Holt, 1954; Klopfers, 1962) was used in 3 of the 14 confirmatory-
validity studies and 3 of the 15 exploratory studies. No other scoring systems were used as frequently. In several of the studies, investigators created their own rules for scoring Rorschach protocols.

The results indicate that validity did not differ across Rorschach scoring systems. For confirmatory studies alone, the results from a Kruskal-Wallis one-way ANOVA were not statistically significant when Exner’s Comprehensive System was compared with all other scoring systems, $\chi^2(1, N = 14) = 0.21, p > .05$, or when comparisons were made among the Comprehensive System, Klopfer’s scoring system, and the remaining scoring systems, $\chi^2(2, N = 14) = 0.21, p > .05$. Results were also not statistically significant when the confirmatory and exploratory studies were combined. This was true when Exner’s Comprehensive System was compared with all other scoring systems, $\chi^2(1, N = 29) = 0.16, p > .05$, and when comparisons were made among Exner’s Comprehensive System, Klopfer’s scoring system, and the remaining scoring systems, $\chi^2(2, N = 29) = 0.45, p > .05$.

Type of statistic

We also analyzed data by type of statistic ($\eta^2$, $f$, $t$, and $\chi^2$). For this meta-analysis, effect-size estimates for the Rorschach and the MMPI from confirmatory and exploratory studies were all pooled together. As before, studies were excluded if there was shared methodological variance between the assessment instruments and the criteria or if effect size was negatively related to validity. When a Kruskal-Wallis one-way ANOVA was conducted, the main effect for type of statistic was not statistically significant, $\chi^2(3, N = 140) = 6.06, p > .10$.

DISCUSSION

The results of the confirmatory studies indicate that validity is better for the MMPI than for the Rorschach. The MMPI explained 23% to 30% of the variance, whereas the Rorschach explained only 8% to 13% of the variance.

In the exploratory studies, both the Rorschach and the MMPI explained only a negligible amount of the variance (3% and 1% for the Rorschach and the MMPI, respectively). However, there was no theoretical or empirical reason to believe that the Rorschach or MMPI would do well. For example, in many of these studies, two or three of the MMPI validity and clinical scales were correlated with a criterion measure. If two or three of the scales are strongly related to the criterion but the other scales are unrelated to the criterion, then the median effect size will be zero. Thus, median effect sizes of zero were often assigned to the MMPI even though the MMPI could be used to make accurate judgments for the task (accurate judgments could be made by attending to only two or three of the scales).

One limitation of the Rorschach studies should be noted. Exner’s Comprehensive System was used in relatively few of these studies, and results may be more promising when Exner’s system is used. However, in our reanalysis of the Parker et al. data, results were not more positive when Exner’s system was used.

The results also indicate that estimates of effect size did not differ significantly depending on whether correlations, $t$ tests, ANOVAs, or chi-square analyses were used to analyze data in the primary articles. This finding undermines the claim by Parker et al. (1988) that correlational statistics are somehow more powerful than the other analyses.

Questions can be raised about the use of the Rorschach in clinical and forensic practice. Our reanalysis of the Parker et al. data indicates that the Rorschach is not as valid as the MMPI. Given findings from other studies on the Rorschach (e.g., that the incremental validity of the Rorschach is poor; Archer & Krishnamurthy, 1997; Garb, 1984, 1998), we recommend that less emphasis be placed on training in the use of the Rorschach. In fact, psychology graduate students may benefit more from a class on judgment and decision making and the use of structured interviews than a class on the Rorschach.

Acknowledgments—We express our appreciation to Kevin C.H. Parker for giving us a list of the studies that were assigned to the convergent-validity and unknown-validity categories and a list of the effect-size estimates for each of the studies.

REFERENCES


(RECEIVED 7/21/97; REVISION ACCEPTED 1/5/98)