

## PROTECTING THE INTEGRITY OF RORSCHACH EXPERT WITNESSES

### A Reply to Grove and Barden (1999) Re: The Admissibility of Testimony Under *Daubert/Kumho* Analyses

Barry Ritzler  
Long Island University

Robert Erard  
Psychological Institutes of Michigan, P.C.

Gary Pettigrew  
Eastern Louisiana Mental Health System

The Rorschach Comprehensive System has been considered by W. M. Grove and R. C. Barden (1999) as inadmissible for expert psychological testimony according to the guidelines from the *Daubert* (1993), *Joiner* (1997), and *Kumho* (1999) decisions. This article refutes W. M. Grove and R. C. Barden's conclusions, arguing that the Rorschach Comprehensive System is (a) testable, (b) valid and reliable, (c) extensively peer reviewed, (d) associated with a reasonable error rate, (e) standardized, (f) accepted by a relevant and substantial scientific community, and (g) appropriate for a wide range of forensic issues. In drawing their negative conclusions, W. M. Grove and R. C. Barden overlooked or minimized a substantial body of empirical data supporting the reliability and validity of the Rorschach Comprehensive System and misinterpreted the language and intent of the Supreme Court decisions.

The Rorschach Comprehensive System (RCS; Exner, 1993) has been considered by Grove and Barden (1999) as inadmissible for expert psychological testimony.<sup>1</sup> The RCS is a specific approach to the Rorschach that was developed by Exner and first published in 1974. Exner developed the system by incorporating procedures of established empirical reliability and validity. It is the most frequently used approach to the Rorschach (Piotrowski & Keller, 1992). The RCS consists of a standardized method of administration and scoring that produces systematically defined variables for use in interpretation.

It has been proposed by Grove and Barden (1999) that the RCS does not qualify as sufficiently "relevant and reliable" (p. 225) for admission in expert testimony. In support of their conclusion, they cite U.S. Supreme Court rulings in

---

Barry Ritzler, Psychology Department, Long Island University; Robert Erard, Psychological Institutes of Michigan, P.C.; Gary Pettigrew, Eastern Louisiana Mental Health System.

Correspondence concerning this article should be addressed to Barry Ritzler, Psychology Department, Long Island University, University Plaza, Brooklyn, New York 11201. E-mail: barry.ritzler@liu.edu

---

<sup>1</sup> The RCS currently is the most frequently used approach to Rorschach interpretation. Other approaches exist and are occasionally used by expert psychological witnesses, but they differ in terms of administration and interpretation methods from the RCS. Because Grove and Barden (1999) directed their argument specifically toward the RCS, this approach is the sole method under consideration in this reply.

the following three cases that address the admissibility of such testimony in federal courts and in states that have adopted the *Federal Rules of Evidence* (1992): *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993); *General Electric Co. v. Joiner* (1997); and *Kumho Tire Co., Ltd. v. Carmichael* (1999). These cases are often referred to as the *Daubert trilogy*. In particular, Grove and Barden (1999) have claimed that “six factors of scientific analysis or indicia of testimonial reliability can be distinguished in *Daubert* that [were later substantiated in *Joiner* and *Kumho*]” (p. 226).<sup>2</sup>

In *Daubert*, the wording is consistent with Grove and Barden’s (1999) six indicia (and one they add later), although the justices include a caveat stating that “definitive checklist or test does not exist in making preliminary assessment of whether reasoning or methodology underlying expert testimony is scientifically valid and whether that reasoning or methodology properly can be applied to facts in issue” (*Daubert*, 1993, p. 2788).

The Court’s subsequent *Kumho* decision extended *Daubert’s* (1993) application to expert testimony based on “technical or other specialized knowledge” (*Kumho*, 1999, p. 226). Thus, the admissibility of “experience-based” or clinical knowledge underlying an expert’s testimony must also be subjected to the trial judge’s discretion as to its reliability. *Kumho* also reemphasized that the test of evidentiary reliability (or trustworthiness) is “flexible.” Justice Breyer, writing for the majority, illustrates this by noting, “it might not be surprising in a particular case, for example, that a claim made by a scientific witness has never been the subject of peer review, for the particular application at issue may never previously have interested any scientist” (*Kumho*, 1999, p. 226). Moreover, a trial court should consider the specific factors identified in *Daubert* only “where there are reasonable measures of the reliability of expert testimony” (*Kumho*, 1999, p. 226).

The seven indicia paraphrased by Grove and Barden (1999) are listed below. These *factors* (in *Daubert/Kumho* terms) are essentially guidelines that judges may (or may not) use in determining admissibility of expert testimony. Neither the original opinion written by Justice Blackmun nor the discussion by Grove and Barden provides specific criteria for testing the admissibility of expert testimony. In applying their seven indicia, however, Grove and Barden have implied or made explicit several specific criteria for judging the RCS. Although we have no argument with the potentially useful guidelines for judging admissibility of expert testimony, we do call into question much of Grove and Barden’s application of their specific criteria and the conclusions drawn from these applications.

The seven *Daubert* guidelines are as follows:

- 1) Is the proposed theory (or technique), on which the testimony is to be based, testable?
- 2) Has the proposed theory (or technique) been tested using valid and reliable procedures and with positive results?
- 3) Has the theory (or technique) been subjected to peer review?

---

<sup>2</sup> The majority opinion of the United States Supreme Court in the *Daubert* case was written by Justice Blackmun, joined by Justices White, O’Connor, Scalia, Kennedy, Souter, and Thomas. Chief Justice Rehnquist, joined by Justice Stevens, filed an opinion concurring in part and dissenting in part.

- 4) What is the known or potential error rate of the scientific theory or technique?
- 5) What standards, controlling the technique's operation, maximize its validity?
- 6) Has the theory (or technique) been generally accepted as valid by a relevant scientific community? (Grove & Barden, 1999, p. 226)
- 7) [Added later] Do the expert's conclusions reasonably follow from applying the theory (or technique) to this case? (Grove & Barden, 1999, p. 226)

In applying these guidelines, Grove and Barden (1999) ultimately have concluded that the RCS is a method that is not admissible as a basis for expert testimony. They have asserted that the RCS meets the acceptable standards for only Guideline 1, that is, it is testable. They have given qualified negative responses to Guidelines 2, 4, 5, and 6 and an outright, unqualified "no" to Guideline 3, that is, error rate. They correctly have refrained from a specific evaluation of Guideline 7, because that requires knowledge of the particular issue involved in the legal case but have used criticism of a review by McCann (1998) to suggest that the RCS fails on this item as well. The general tone of Grove and Barden's conclusions, therefore, is that the RCS can hardly be expected to be appropriate for many court issues.

If taken seriously by judges trying to determine the admissibility of the RCS, Grove and Barden's (1999) conclusions would take the technique out of the hands of expert witnesses in numerous applications for which it is used in thousands of courtrooms each year. Consequently, a reply is necessary from psychologists who consider the RCS to be an appropriate method for expert psychological testimony.

The reply focuses in detail on each of Grove and Barden's (1999) arguments in the hope of providing judges with an alternative perspective for evaluating the *Daubert* guidelines. Note that this reply refers only to the Comprehensive System approach to the Rorschach because (a) Grove and Barden's arguments are directed specifically to the use of the Comprehensive System approach and (b) the Comprehensive System, compared with other Rorschach approaches, has an extensive body of supporting literature.

### Is the RCS Testable?

Grove and Barden (1999) have acknowledged that because the RCS yields a myriad of scores that can easily be tested in regard to external criteria, the RCS is, indeed, testable. This is the sole guideline for which they have granted credibility for the use of the RCS in expert testimony. We, of course, concur with this conclusion.

### Has the RCS Been Tested by Valid and Reliable Procedures With Positive Results?

Grove and Barden (1999) have acknowledged that the RCS has been appropriately tested, but they are equivocal in regard to the "positive" nature of the results. They have observed that in zero-order correlational studies, the RCS has shown what they consider to be "low" validity, "somewhat lower than the [Minnesota Multiphasic Personality Inventory] MMPI." (Grove & Barden, 1999,

p. 227). In support of this contention, they have cited only two of their own reviews: Garb, Florio, and Grove (1998) and Garb, Wood, Nezworski, Grove, and Stejskal (2001). No reference has been made to numerous existing reviews and studies that place the RCS at a more robust level of zero-order validity (see Viglione, 1999). Citing data reported by Archer and Gordon (1988), Meyer (2000) recently addressed this issue with regard to the use of the (RCS Schizophrenia Index; SCZI) for classification of psychotic diagnoses.<sup>3</sup> He noted that the optimal overall hit rate for the SCZI (using a cutoff of 5) was .80 whereas that for the MMPI Scale 8 was .76 (using a cutoff of  $T > 75$ ). When traditional cutoffs were used (SCZI  $> 4$ ), this RCS index outperformed MMPI Scale 8 with a hit rate of .69 as opposed to .48 (for  $T$  score  $> 65$ ) and .60 (for  $T$  score  $> 70$ ).

Grove and Barden (1999) have gone on to state that meta-analyses of the RCS suggest typical RCS score validities of about .30. Actually, such validity estimates have ranged from .30 (Hiller, Rosenthal, Bornstein, Berry, & Brunell-Neuleib, 1999) to .41 (Parker, Hanson, & Hunsley, 1988), with all mean validity figures in the moderate range compared with other psychological measures (Fiske & Campbell, 1992). Note that all personality tests produce validity coefficients in the moderate range, unless one intervenes to align method variance artificially (e.g., correlate one self-report scale with another; use items from a self-report scale that are correlated with the same questions asked of the client during a structured interview; or use scales composed of unrealistic virtues or problems to detect those who were encouraged to report unrealistic virtues or problems in a fake-good or fake-bad paradigm.) Most meta-analytic studies have found no significant difference between the RCS and the MMPI; the latter being generally accepted among psychologists and the courts as representing the state of the art in clinical personality assessment (Archer, 1992; Nichols, 1992; Ogloff, 1995; and Pope, Butcher, & Seelen, 1993). Grove and Barden have disparaged a .30 mean validity by analyzing the error rate of a "Gaussian measure" with .30 validity. In fact, many RCS measures and clinical personality traits are not Gaussian, that is, normally distributed. In contrast, Hiller et al., using a more appropriate form of analysis, have shown that at a .30 level, a hypothetical dichotomous RCS indicator would support some theoretically relevant dichotomous criterion 64.5% of the time.

Grove and Barden's (1999) only allusion to literature supporting RCS validity is the following comment: "Rorschach proponents see the validity more optimistically than do skeptics" (p. 227). Apart from providing an operational definition of *proponents* and *skeptics*, this comment throws little light on the admissibility of the RCS for expert testimony.

Next, Grove and Barden (1999) have admitted, "One could argue that if the Rorschach has any zero-order validity, it passes this *Daubert* test" (p. 227). However, they quickly have retracted this qualified support of the RCS by stating that the critical test of positive results is with incremental validity concept found nowhere in the *Daubert* opinion or its progeny, but which turns out to be a

---

<sup>3</sup> The SCZI recently has been modified to the Perceptual Thinking Index—a modification that has resulted in an even more valid indicator of cognitive dysfunction (J. E. Exner, personal communication, September, 2001).

hobbyhorse of Grove and Barden and other RCS critics. Nevertheless, with regard to this issue, note that the SCZI and the Depression Index (DEPI) are useful in supporting the general diagnoses of psychosis and depression beyond information available from the second edition of the MMPI (MMPI-2).<sup>4</sup> Meyer (2000) found that the SCZI made a substantial and unique contribution to the prediction of psychotic disorder diagnosis and the DEPI offered a modest contribution to the prediction of depressive disorders over that which could be obtained from the MMPI-2.

Apart from diagnosis, empirical and logical demonstrations have shown how the RCS frequently has incremental validity above that obtained by self-report methods alone (e.g., Acklin, 1993; Ganellen, 1996a, 1996b; Meyer, 1997, 1999a, 1999b, Meyer & Handler, 2000; Meyer, Riethmiller, Brooks, Benoit, & Handler, 2000; Stricker and Gold, 1999; Viglione, 1996, 1999). *Incremental validity* refers to a method's ability to provide accurate information beyond that of comparison methods. For example, the RCS would have incremental validity in relation to self-report methods if it provided accurate information not obtainable from the self-report methods used at the same time with the same individuals. Meyer (1999a) has demonstrated that patterns obtained from a combined use of the RCS and the MMPI (i.e., "cross-method patterns") were meaningfully related to diagnoses at a higher level than either method alone. In contrast, Grove and Barden (1999) have cited only a single review of Rorschach incremental validity written by one of their colleagues (Garb, 1985), which was completed years before the studies cited above and included no RCS studies.

Addressing the larger issue of scientific testimony of any sort, Justice Blackmun stated that "scientific validity for one purpose is not necessarily scientific validity for other, unrelated purposes" (Daubert, 1993, p. 2788). In arguing that statistically demonstrated incremental validity constitutes the sine qua non of forensic utility, Grove and Barden (1999) ignored the principle of validity for specific purposes. Hiller et al. (1999) found that the MMPI has larger validity coefficients than the RCS for studies using psychiatric diagnoses and self-report measures as criterion variables; whereas the RCS had larger validity coefficients than the MMPI for studies using objective criterion variables.

After their discussion of the second guideline, Grove and Barden (1999) have labeled as *Guideline 3* the question, "Is the Rorschach generally accepted as valid for diagnosis and personality description?" in place of the original Guideline 3, which asked, "Has the theory (technique) been subjected to peer review?" This original Guideline 3 is addressed later in the discussion and is renumbered as Guideline 5. To overcome this confusion, the new question is addressed first and the original question is returned to immediately afterward.

### Is the Rorschach Generally Accepted as Valid for Diagnosis and Personality Description?

Grove and Barden (1999) have refuted emphatically the RCS's status as a generally accepted method while citing only the Buros's *Thirteenth Mental*

---

<sup>4</sup> The SCZI and the DEPI are two indices that consist of a combination of RCS variables that are associated with the presence of schizophrenic and depressive symptoms.

*Measurements Yearbook* (Impara & Plake, 1988), two outdated reviews (Cronbach, 1949, and Jensen, 1965), and two reviews from their previously mentioned circle of colleagues (Nezworski & Wood, 1995, and Wood, Nezworski, & Stejskal, 1996). It seems preferable to separate the new question into its two components: Is the RCS generally accepted as valid for (a) diagnosis and (b) personality description? With regard to its validity for diagnosis, using the *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)* (American Psychiatric Association, 1994) as a criterion, the RCS is, indeed, not very valid. The RCS does not significantly correlate with many *DSM-IV* categories (although there is a relationship to general psychotic and depressive classifications). But what of the validity of the *DSM-IV*? Because Grove and Barden deplore the validity of the RCS, they must surely have to admit that *DSM-IV* is by no means a purely empirical classification system. Indeed, in its introduction, its developers characterize the *DSM-IV* as a system for categorization of psychiatric syndromes developed by committee and stress that it is not a measure of pathology. Admittedly, courts tend to perceive the *DSM-IV* as a highly valid and comprehensive indicator of psychopathology; a sort of gold standard for expert testimony (McCann & Dyer, 1996). Nevertheless, even the American Psychiatric Association (1994) has warned, "When the *DSM-IV* categories, criteria, and textual descriptions are employed for forensic purposes, there are significant risks that diagnostic information will be misused or misunderstood" (p. xxiii). Later in their article, as if agreeing with this point, Grove and Barden judge dissociative identity disorder (multiple personality), a *DSM-IV* category, as inadmissible for use in expert testimony. Furthermore, we argue that the RCS provides a more useful personality description than the *DSM-IV*, which tends to include very different individuals in the same diagnostic category. Diagnosis is at best the beginning of the study of a person's psychological functioning, not its endpoint (Persons, 1986). For instance, the *DSM-IV* may diagnose as schizophrenic two persons, one socially active and loquacious and another withdrawn and silent. The RCS can easily distinguish between these two different persons while identifying their schizophrenic characteristics. For instance, although both persons are likely to score significantly on the SCZI, the socially active patient is likely to give more responses with human content, more complicated responses, and more responses indicative of overt expressions of psychosis compared with the responses of the withdrawn patient.

The RCS literature is replete with studies showing a positive correspondence between RCS variables and personality characteristics and behaviors (e.g., Bornstein, Hill, Robinson, Calabrese, & Bowers, 1996, on dependency<sup>5</sup>; Burns & Viglione, 1996, 1997, on the RCS Human Experience variable; Meyer, 1999b, on the RCS Prognostic Rating Scale<sup>6</sup>; Weiner, 1996, on the validity of the RCS; Weiner, 1999, on treatment planning and outcome assessment; and Gacono & Meloy, 1994, on aggressive and psychopathic personalities). This does not include the large body of literature that was used originally by Exner (1993) to develop

---

<sup>5</sup> This dependency variable currently is being evaluated by the RCS research council for possible inclusion in the RCS.

<sup>6</sup> The Prognostic Rating Scale is not an RCS index, but its component variables closely correspond to current RCS variables.

the RCS or the many studies that he and his colleagues conducted in establishing nationally representative norms and testing relationships posited by the earlier Rorschach literature. With over 200 books and between 8,000 and 9,000 articles on the subject, the Rorschach is the second most thoroughly researched personality instrument in history (Butcher & Rouse, 1996; Exner, 1997), and with nearly 3,000 members in the Society for Personality Assessment (most of whom frequently use the RCS in their practice) and even more in the International Rorschach Society, Rorschach experts can certainly provide an adequate sampling of learned treatises, journal articles, and more than a recognized minority of clinical scientists to satisfy the courts' concerns. Grove and Barden (1999) appear to have disregarded these hundreds of studies and individual proponents. Simply stated, the answer to this new Question 3 hinges on Guideline 7; that is, whether the technique is appropriate for the issue being addressed. If the issue is the *DSM-IV* diagnosis, an opinion based on the RCS alone may not be admissible. Actually, we are in favor of excluding expert testimony that attempts to use the RCS *in isolation* to determine the correct diagnostic label for a person. If, on the other hand, personality description is relevant to the issue in question, the RCS can be quite useful and can meaningfully supplement the basis for testimony derived from other techniques, including the *DSM-IV*.

### Has the Theory Been Subjected to Peer Review?

This question, originally numbered 3 by Grove and Barden (1999), was given as Number 5 in their discussion. Their answer was yes, no, and maybe. In trying to explain their omnibus conclusion, they began by stating that "many of Exner's studies cited in his 1993 volume as supporting [RCS] were apparently never peer reviewed; in fact, many do not even exist as actual reports, and their data are not readily available to other investigators." (p. 228). It is true that some of the studies cited as empirical support for the RCS are unpublished manuscripts—usually dissertations and other student research projects performed with Exner as mentor. Grove and Barden have not specified what they mean by "many," but an actual count taken from Exner's (1993) text reveals that fewer than 10% of the supporting studies were cited as "unpublished manuscripts." This indicates that over 400 provided empirical support for the RCS *were* published in peer-reviewed journals.<sup>7</sup> Furthermore, in a recent survey of manuscripts concerning personality assessment methods appearing in peer-reviewed journals over a 10-year span (1985–1995), the Rorschach was a close second to the MMPI-A (Minnesota Multiphasic Personality Inventory for Adolescents) in the number of manuscripts published that supported its validity and/or reliability (Ritzler, 1996).

Grove and Barden (1999) go on to state that "pro-Rorschach studies often appear in a specialty journal, the *Journal of Personality Assessment*. . . which may potentially relate to quality of publication" (p. 228). This statement may cause considerable dismay for the editors, reviewers, and contributors to the *Journal of*

---

<sup>7</sup> According to the Ninth Circuit, further interpreting the Supreme Court standards, if an expert's theories and methods have been subjected to peer review and accepted by even a small portion of the scientific community, the expert's opinion may be based on "scientifically valid principles."

*Personality Assessment*, a journal with a 60-year history of peer-reviewed publications with current rejection rates running between 80% and 85%. Most of the professional journals in psychology could be considered specialty journals (e.g., *Journal of Abnormal Psychology* or *Psychology, Public Policy, and Law*). To date, no evidence has been presented to suggest that the peer reviews in such journals are inferior to the reviews in more general publications (e.g., *American Psychologist* or *Journal of Clinical Psychology*). In fact, articles supporting the validity of the RCS and other systems for Rorschach interpretation have appeared in journals other than the *Journal of Personality Assessment*. Consequently, the RCS surely qualifies for a definite yes on Guideline 3.

### What Is the Known or Potential Error Rate of the RCS?

From a legal standpoint, *error rate* ordinarily refers to the probability that the application of a particular procedure or theory can lead to a mistake in the classification or identification of an object, event, or person (Mauet & Wolfson, 1997). Nowhere in the *Daubert* trilogy is there a formula for deriving an error rate or a standard beneath which the error rate may not fall. Furthermore, it is neither possible nor reasonable to apply the concept of error rate to the RCS as a whole. Indeed, it is inappropriate to speak globally of the “validity of the test” (Rogers, Salekin, & Sewell, 1999). By analogy, let us consider the question, “What is the error rate of X-rays as used in medicine?” Certainly, X-rays find their way into the courtroom almost every day with few challenges under *Daubert*, despite the fact that it is impossible to assign a global figure to overall validity as a technique for imaging the human body. If we are referring to X-rays of bony fractures, do we mean X-rays of gross fractures or hairline fractures? Are we using mammography to locate possible tumors, and are we looking at tumors large enough to be palpable on manual examination or extremely small tumors not likely to be found on manual examination? Are we using new, state-of-the-art imaging equipment with interpretation by board-certified radiologists or old, outdated equipment with interpretation by internists with little experience in radiology? Of course, it is possible in theory to establish the error rate of a particular type of X-ray under minutely specified conditions assuming that a surgical or other type of gold standard is available. But what benchmark can we use in personality assessment?

When speaking about validity, we make a distinction between criterion validity (as when specific diagnoses or behavioral predictions are made—e.g., the patient is or is not depressed, will or will not respond to a certain treatment, or will or will not commit suicide) and construct validity (i.e., the extent to which a construct or theory has been validated through different experiments under different conditions with different samples to form a more integrated and generalizable understanding of the construct or theory). Although the RCS is not primarily used for either diagnosis or behavioral predication in isolation, Grove and Barden (1999) seem to have concerned themselves almost exclusively with questions of criterion validity. In those few areas in which RCS findings have been linked to specific diagnostic and behavioral criteria (e.g., the SCZI, the Rorschach Prognostic Rating Scale, the DEPI, and the Suicide Constellation), known error rates have been established (cf. Hilsenroth, Fowler, & Padawer, 1998; Meyer & Handler, 2000; Viglione, 2000). Far more important, however, is



the use of the RCS for personality description, where construct validity more properly applies. As Weiner (2000) has observed, “The validity of Rorschach findings is appropriately assessed only by examining their correlates with conditions and behaviors that are conceptually linked by dimensions of personality functioning to the Rorschach variables being examined (p. 167).” This is construct validity, not the simpler and less informative criterion validity.

Appropriate procedures for establishing construct validity are described in the *Standards for Educational and Psychological Testing—Revised* (American Psychological Association, Committee on Professional Standards, 2000). Construct validity involves the use of multiple heterogeneous measures, each of which either theoretically reflects the construct in some fashion or is theoretically incompatible with the construct and is used to establish convergent and discriminant validity. As construct validity develops, the validity of a measure of that construct is best understood in terms of *goodness of fit* rather than *error rate*. For example, when examining a portrait, the questions becomes not, “What is the likelihood that this is a portrait of Whistler’s mother?” but rather, “To what degree and by what standard can we call this painting a good likeness of her?”

The evaluation of error rates for construct validity is much more complicated than the single-statistic assessment of criterion validity. One can approach this problem with the RCS by applying decision theory to the mean *effect sizes* of a host of studies comparing RCS variables to various theoretically linked observations. An effect size is the mathematical magnitude of difference between significantly different variables. When a sufficient number of effect sizes are obtained from numerous validation studies of a particular method, error rates can be calculated. A review of such effect sizes for both the RCS and the MMPI shows that the two methods are quite comparable in terms of overall error rates (Hiller et al., 1999). Even so, error rates are much more appropriate for estimating the criterion validity of a variable rather than the construct validity of a more comprehensive method (i.e., one that involves a system of many variables). Consequently, we would not advise courts to rely too heavily on the concept of error rate as a test for the admissibility of any measure of a psychological construct. A more appropriate gauge of admissibility would be the extent to which the current scientific literature establishes that a particular method for measuring a construct works as expected and as well or better than any other in providing a good fit for that construct, that is, construct validity (cf. Meyer, 1996; Meyer et al., 2000). The primary use of the RCS is not to predict specific behaviors, but instead to help clinicians “describe the complex interaction among psychological, biological, environmental, and behavioral domains” (Viglione, 2000, p. 260)—exactly the kind of interaction with which many important forensic questions concern themselves. For instance, the RCS is not very good at predicting that a person will commit a certain kind of crime, but once that crime is committed, the RCS is very good at enabling and understanding of the personality characteristics associated with the crime and can be particularly helpful in planning rehabilitation, assigning appropriate sentencing, and/or identifying mitigating or extenuating factors.

Grove and Barden (1999) have suggested that it would be reasonable to use “one or a few well-validated scores” (p. 229) from the RCS, but they go on to criticize those who attempt to integrate Rorschach data with interview information, life history data, and other psychological test scores on the grounds that such

integration may “wash out” valid diagnostic information and contaminate inferences in a way that may obscure error rates. This approach by Grove and Barden indicates how very far they have wandered from the mainstream of clinical psychology. It is precisely because there are no perfectly valid personality measures and because no test is self-interpreting (Meyer & Handler, 2000) that the integration of multiple methods is necessary in assessment in general and in forensic evaluations in particular. Grove and Barden’s efforts to eliminate the expert from expert testimony by insisting that clinical measures speak for themselves and in isolation from their clinical context are out of step with the multimethod approach that lies at the heart of good personality assessment (American Psychological Association, 2000; Anastasi & Urbina, 1997; Dyer & McCann, 2000; and Matarazzo, 1991, 1992). *Daubert* does not require that the expert provide an estimate of the isolated probabilities regarding the correctness of each of his or her opinions. Instead, it recommends that there is a reasonable effort to find appropriate means of testing the validity of the methodology underlying those opinions.

#### What Standards, Controlling the Technique’s Operation, Maximize Its Validity?

Curiously, Grove and Barden have not addressed this guideline in their discussion. This may be because the RCS rates well in this area. There are specific instructions for administration and coding (Exner, 1995) that set high standards for correct procedure in using the technique. Some psychologists do not adhere to these standards when using the RCS, and as a result, their testimony would not qualify as expert under this guideline. However, the RCS certainly provides the standards of operations necessary for maximizing the technique’s validity.

#### Has the RCS Been Generally Accepted as Valid in the Relevant Scientific Community?

The answer to this question is yes. The RCS and other approaches to the Rorschach are taught in over 85% of graduate programs in psychology accredited by the American Psychological Association. Furthermore, a survey of 412 clinical psychologists engaged in providing assessment services (Watkins, Campbell, Nieberding, & Hall, 1995) found that 82% of respondents used the Rorschach in their assessments. A subsequent American Psychological Association survey (Camara, Nathan, & Puente, 1998) confirmed that the Rorschach is one of the most frequently used psychological tests. Also, Archer and Newsom (2000) recently obtained updated survey information from 346 psychologists who spend a significant amount of time working with adolescents. The Rorschach was the second most frequently used test; second only to the Wechsler intelligence scales and considerably more frequently used than the MMPI–A and the Millon Adolescent Clinical Inventory. Furthermore, at least two relatively large professional organizations support the continuing use and study of the technique: the International Rorschach Society, with a worldwide membership, and the Society for Personality Assessment, with nearly 3,000 members from the United States and other countries. Certainly, the technique has its skeptics and detractors, as does the

MMPI, *DSM-IV*, and other widely accepted procedures, but the opinions of 5,000 to 10,000 psychologists demonstrate that it is a generally accepted technique.

### Conclusion

When the criteria outlined in *Daubert* and its progeny are applied, the RCS meets the requirements for admissibility. Our disagreement with Grove and Barden (1999), however, is not limited to this single issue. Our disagreement extends to a major difference in philosophy about how clinical assessment should proceed and its overall place in expert testimony.

In their 1999 article, Grove and Barden ultimately argue that their target is not merely psychological expert testimony based on the use of the RCS, but rather all projective testing and, consequently, all psychological expert testimony based on any procedure or technique that derives from clinical judgment rather than statistical prediction. They argue that they are protecting “the integrity of the legal system” (p. 238) from misleading and irresponsible expert testimony. Most of their arguments against the RCS apply equally well to the use of the MMPI-2, the Millon tests, the *DSM-IV*, clinical interviews, mental status examinations, and, ultimately, any application of clinical psychological principles. They seem to regard the well-established multimethod, multisource approach to developing convergent validity (American Psychological Association, Assessment Work Group, 1998; American Psychological Association, Committee on Professional Standards, 2000; Dyer & McCann, 2000; Heilbrun, 1992; Matarazzo, 1991, 1992; and Wiggins, 1973) as merely a means of generating actuarial uncertainty. Thus, rather than relying on the highest standards of the relevant professional community, as the *Daubert* court requires, Grove and Barden would paradoxically limit our permissible testimony to simple truisms that prevent effective application of expert knowledge and experience.

In fact, the relevant empirical research on the impact of expert testimony on juries and other finders of fact casts serious doubt on whether the legal system needs the kind of ultrastringent protection Grove and Barden (1999) advocate. Voir dire, cross-examination, the rules of evidence, judges’ instructions, and jurors’ reliance on the totality of the evidence and their common sense routinely prevent seduction by “experts” armed with dubious theories and techniques (cf. Vidmar et al., 2000). Attempting to exclude, in effect, all expert testimony that is not established beyond the reach of all scientific controversy would paradoxically prevent legal fact finders from gaining access to the raw materials required to do their job.<sup>8</sup> In complex litigation, jurors need access to professional opinions from fields such as psychology, medicine, economics, and engineering, which are

---

<sup>8</sup> In the Ninth Circuit’s decision on remand in *Daubert*, which is heavily relied on by legal experts for interpretation of *Daubert*’s legal significance (McDonald, 2000), Judge Alex Kozinski concluded, “that the testimony proffered by an expert is based directly on legitimate, preexisting research unrelated to the litigation provides the most persuasive basis for concluding that the opinions he expresses were derived by the scientific method” (p. 1307). Judge Kozinski further commented that experts can “explain precisely how they went about reaching their conclusions and point to some objective source, a published article in a reputable scientific journal or the like to show that they have followed the scientific method, as it is practiced by (at least) a recognized minority of scientists in their field” (p. 1319).

rarely based on the kind of simple predictive relationships favored by Grove and Barden, but rather on a complex blend of theory, research, and professional judgment.

Grove and Barden (1999) are correct in their observation that the *Daubert* decision broadens the scrutiny to which psychological experts will be subjected. Even so, the RCS provides a model for the orderly evolution of the kind of scientifically based, state-of-the-art clinical practice that *Daubert* challenges us to uphold.

## References

- Acklin, M. W. (1993). Integrating the Rorschach and the MMPI in clinical assessment: Conceptual and methodological issues. *Journal of Personality Assessment*, *60*, 125–131.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- American Psychological Association, Assessment Workgroup. (1998). *Benefits and costs of psychological assessment in healthcare delivery: Report of the Board of Professional Affairs, Part 1*. Washington, DC: Author.
- American Psychological Association, Committee on Professional Standards. (2000). *Standards for educational and psychological testing—Revised*. Washington, DC: Author.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Archer, R. P. (1992). Minnesota Multiphasic Personality Inventory—2. In J. J. Kramer & J. C. Conoley (Eds.), *Eleventh mental measurements yearbook* (pp. 558–562). Lincoln, NE: Buros Institute of Mental Measurements.
- Archer, R. P., & Gordon, R. A. (1988). MMPI and Rorschach indices of schizophrenic and depressive diagnoses among adolescent inpatients. *Journal of Personality Assessment*, *52*, 276–287.
- Archer, R. P., & Newsom, C. R. (2000). Psychological test usage with adolescent clients: Survey update. *Assessment*, *7*, 227–235.
- Bornstein, R. F., Hill, E. L., Robinson, K. J., Calabrese, C., & Bowers, K. S. (1996). Internal reliability of Rorschach Oral Dependency Scale scores. *Educational and Psychological Measurement*, *56*, 130–138.
- Burns, B., & Viglione, D. J. (1996). The Rorschach Human Experience variable, interpersonal relatedness, and object representations in nonpatients. *Psychological Assessment*, *8*, 92–99.
- Burns, B., & Viglione, D. J. (1997). “The Rorschach Human Experience variable, interpersonal relatedness, and object representations in nonpatients”: Correction. *Psychological Assessment*, *9*, 82.
- Butcher, J. N., & Rouse, S. V. (1996). Personality: Individual differences and clinical assessment. *Annual Review of Psychology*, *47*, 87–111.
- Camara, W. J., Nathan, J. S., & Puente, A. E. (2000). Psychological test usage: Implications in professional psychology. *Professional Psychological Research and Practice*, *31*, 141–154.
- Cronbach, L. J. (1949). “Pattern tabulation”: A statistical method for analysis of limited patterns of scores, with particular reference to the Rorschach test. *Educational and Psychological Measurement*, *9*, 149–171.
- Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U. S., 113 S. Ct. 2786 (1993).
- Dyer, F. J., & McCann, J. T. (2000). The Millon clinical inventories, research critical of their forensic application, and *Daubert* criteria. *Law and Human Behavior*, *24*, 487–497.

- Exner, J. E. (1974). *The Rorschach: A comprehensive system. Vol. 1: Basic foundations*. New York: Wiley.
- Exner, J. E. (1993). *The Rorschach: A comprehensive system. Vol. 1: Basic foundations* (3<sup>rd</sup> ed.). New York: Wiley.
- Exner, J. E. (1995). *A Rorschach workbook for the comprehensive system* (4th ed.). Asheville, NC: Rorschach Workshops.
- Exner, J. E. (1997). The future of Rorschach in personality assessment. *Journal of Personality Assessment*, 68, 37–46.
- Federal rules of evidence*. (1992). Boston: Little, Brown.
- Fiske, D. W., & Campbell, D. T. (1992). Citations do not solve problems. *Psychological Bulletin*, 122, 393–395.
- Gacono, C. B., & Meloy, J. R. (1994). *The Rorschach assessment of aggressive and psychopathic personalities*. Hillsdale, NJ: Erlbaum.
- Ganellen, R. J. (1996a). Comparing the diagnostic efficacy of the MMPI, MCMI–II, and Rorschach: A review. *Journal of Personality Assessment*, 67(2), 219–243.
- Ganellen, R. J. (1996b). Exploring MMPI–Rorschach relationships. *Journal of Personality Assessment*, 67, 529–542.
- Garb, H. N. (1985). Linear and configural statistical models in the diagnosis of mental disorders. *Dissertations Abstracts International*, 45(7B), 2307, University of Minnesota.
- Garb, H. N., Florio, C. M., & Grove, W. M. (1998). The validity of the Rorschach and the Minnesota Multiphasic Personality Inventory: Results from meta-analyses. *Psychological Science*, 9, 402–404.
- Garb, H. N., Wood, J. M., Nezworski, M. T., Grove, W. M., & Stejskal, W. J. (2001). Toward a resolution of the Rorschach controversy. *Psychological Assessment*, 13, 433–448.
- General Electric Co. v. Joiner*, 118 S. Ct. 512 (1997).
- Grove, W. M., & Barden, R. C. (1999). Protecting the integrity of the legal system: The admissibility of testimony from mental health experts under *Daubert/Kumho* analyses. *Psychology, Public Policy, and Law*, 5, 224–242.
- Heilbrun, K. (1992). The role of psychological testing in forensic assessment. *Law and Human Behavior*, 16, 257.
- 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., Fowler, J. C., & Padawer, J. R. (1998). The Rorschach Schizophrenia Index (SCZI): An examination of reliability, validity, and diagnostic efficiency. *Journal of Personality Assessment*, 70, 514–534.
- Impara, J. C., & Plake, B. S. (Eds.). (1988). *Thirteenth mental measurements yearbook*. Lincoln, NE: Buros Institute.
- 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.
- Kumho Tire Co., Ltd. v. Carmichael*, 119 S. Ct. 1167 (1999).
- Matarazzo, J. D. (1991). Psychological testing is reliable and valid: Reply to Ziskin and Faust. *American Psychologist*, 46, 882–884.
- Matarazzo, J. D. (1992). Psychological testing and assessment in the 21st century. *American Psychologist*, 47, 1007–1018.
- Mauet, T. A., & Wolfson, W. D. (1997). *Trial evidence*. New York: Aspen Law & Business.
- McCann, J. T. (1998). Defending the Rorschach in court: An analysis of admissibility using legal and professional standards. *Journal of Personality Assessment*, 70, 125–144.

- McCann, J. T., & Dyer, F. J. (1996). *Forensic assessment with the Millon inventories*. New York: Guilford Press.
- McDonald, J. J. (2000). *Mental and emotional injuries in employment litigation: 2000 cumulative supplement*. Washington, DC: U.S. Bureau of National Affairs
- Meyer, G. (1996). The Rorschach and the MMPI: Toward a more scientifically differentiated understanding of cross-method assessment. *Journal of Personality Assessment*, *67*, 558–578.
- Meyer, G. J. (1997). Thinking clearly about reliability: More critical corrections regarding the Rorschach Comprehensive System. *Psychological Assessment*, *9*, 495–498.
- Meyer, G. J. (1999a). 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.
- Meyer, G. J. (1999b). Introduction to the first special section in the special series on the utility of the Rorschach for clinical assessment. *Psychological Assessment*, *11*, 235–239.
- Meyer, G. J. (2000). On the science of Rorschach research. *Journal of Personality Assessment*, *75*, 46–81.
- Meyer, G. J., & Handler, L. (2000). Incremental validity of the Rorschach Prognostic Rating scale over the MMPI Ego Strength scale and IQ. *Journal of Personality Assessment*, *74* (3), 356–370.
- Meyer, G. J., Riethmiller, R. J., Brooks R. D., Benoit, W. A., & Handler, L. (2000). A replication of Rorschach and MMPI–2 convergent validity. *Journal of Personality Assessment*, *74*, 175–215.
- Nezworski, M. T., & Wood, J. M. (1995). Narcissism in the Comprehensive System for the Rorschach. *Clinical Psychology: Science and Practice*, *2*(2), 179–199.
- Nichols, D. S. (1992) Minnesota Multiphasic Personality Inventory–2. In J. J. Kramer & J. C. Conoley (Eds.), *Eleventh mental measurements yearbook* (pp. 562–565). Lincoln, NE: Buros Institute of Mental Measurements.
- Ogloff, J. R. P. (1995). The legal basis of forensic applications of the MMPI–2. In Y. S. Ben-Porath, J. R. Graham, G. C. N. Hall, R. D. Hirschman, & M. S. Zaragoza (Eds.), *Forensic applications of the MMPI–2* (pp. 18–47). Thousand Oaks, CA: Sage.
- Parker, K. C., Hanson R. K., & Hunsley, J. (1988). MMPI, Rorschach, and WAIS: A meta-analytic comparison of reliability, stability and validity. *Psychological Bulletin*, *103*, 367–373.
- Persons, J. B. (1986). The advantages of studying psychological phenomena rather than psychiatric diagnoses. *American Psychologist*, 1252–1260.
- Piotrowski, C., & Keller, J. (1992). Psychological testing in applied settings: A literature review from 1982–1992. *Journal of Training and Practice in Professional Psychology*, *6*, 74–82.
- Pope, K. S., Butcher, J. N., & Seelen, J. (1993). *The MMPI, MMPI–2, and MMPI–A in court: A practical guide for expert witnesses and attorneys*. Washington, DC: American Psychological Association.
- Ritzler, B. (1996). *The status of personality assessment*. Presidential address presented at the meeting of the Society for Personality Assessment, Denver, CO.
- Rogers, R., Salekin, R., & Sewell, K. (1999). Validation of the Millon Clinical Multiaxial Inventory for Axis II disorders: Does it meet the *Daubert* standard? *Law and Human Behavior*, *23*, 425–443.
- Stricker, G., & Gold, J. R. (1999). The Rorschach: Toward a nomothetically based, idiographically applicable configurational model. *Psychological Assessment*, *11*, 240–250.
- Vidmar, N., Lempert, R. O., Diamond, S., Hans, V. P., Landsman, S., MacCoun, R., et al.

- (2000). Amicus brief: *Kumho Tire v. Carmichael*. *Law and Human Behavior*, 24, 387-400.
- Viglione, D. J. (1996). Data and issues to consider in reconciling self-report and the Rorschach. *Journal of Personality Assessment*, 67, 579-587.
- Viglione, D. J. (1999). A review of recent research addressing the utility of the Rorschach. *Psychological Assessment*, 11, 251-265.
- Viglione, D. J. (2000). On the science of Rorschach research. *Journal of Personality Assessment*, 75, 46-81.
- Watkins, C. E., Campbell, V. L., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychological Research and Practice*, 26, 54-60.
- Weiner, I. B. (1996). Some observations on the validity of the Rorschach inkblot method. *Psychological Assessment*, 8, 206-213.
- Weiner, I. B. (1999). Contemporary perspectives on Rorschach assessment. *European Journal of Psychological Assessment*, 15(1), 78-86.
- Weiner, I. B. (2000). Making Rorschach interpretation as good as it can be. *Journal of Personality Assessment*, 74, 164-174.
- Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*. Reading, MA: Addison-Wesley.
- Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1996). The comprehensive system for the Rorschach: A critical examination. *Psychological Science*, 7, 3-10.
- Wood, J. M., Nezworski, M. T., Stejskal, W. J., Garven, S., & West, S. G. (1999). "Methodological issues in evaluating Rorschach validity: A comment on Burns and Viglione (1996), Weiner (1996), and Ganellen (1996)": Erratum. *Assessment*, 6, 305.