JOURNAL OF PERSONALITY ASSESSMENT, *83*(2), 141–152

ARTICLES

A Consideration of Challenges to Psychological Assessment Instruments Used in Forensic Settings: Rorschach As Exemplar

Mark J. Hilsenroth and George Stricker
The Derner Institute of Advanced Psychological Studies
Adelphi University

In this brief primer, we provide an outline of key issues that will help psychologists organize and prepare their expert testimony. These issues include the need to obtain essential sources of research, a review of the actual legal standards regarding admissibility of test data in expert testimony, the nature of the expert relative to the assessment instrument in expert testimony, the nature of legal versus scientific debate, and the examination of appropriate qualifications of expertise when offering legal testimony. In addition, we use a summary of information contained in several recent articles to address challenges directed against forensic psychological testing. We use the empirical literature on the Rorschach as an exemplar in discussing these issues, as the admissibility of the Rorschach in particular has been challenged, and the issues frequently focused on with the Rorschach are equally applicable to other psychological measures. In this article, we provide essential sources of Rorschach research regarding several empirical studies that summarize important information and directly address previous criticisms of the measure.

In this brief primer, we provide a summary of how the information contained in several recent articles might be used to address challenges to forensic psychological testing. Forensic psychologists must be acquainted with the arguments of critics, their strengths, and their vulnerabilities to prepare properly for any testimony involving psychological testing. We use the empirical literature on the Rorschach as an exemplar in discussing each of these issues, as the admissibility of the Rorschach in particular has been challenged, and the issues frequently focused on by critics of the Rorschach are equally applicable to other psychological measures.

OBTAIN ESSENTIAL REFERENCE MATERIALS (REVIEWS AND ORIGINAL SOURCES)

The best preparation for discussing the strengths and limitations of any test is knowing the current research literature on that measure. A close appraisal of recent criticism of a given measure and subsequent response to those criticisms may provide invaluable information regarding strategy within a forensic context. Such knowledge will help to inoculate the jury to these arguments and to bolster the psychologist's response to cross-examination. It also will aid counsel in refutation of opposing expert witnesses on cross-examination. Essential reference sources for any given test include information on reliability, validity, incremental validity, normative data, ethnic diversity issues, specific applications, patterns of use, and clinical as well as forensic utility. Familiarization with such literature is crucial in preparing for potential areas of cross-examination. In addition, the close review of original sources that have been singled out for particular praise or criticism can be highly informative and useful regarding debate. On cross-examination, psychologists may be asked whether they are familiar with important references and would do well to be aware of them. However, from a practical standpoint, it is important not to accept any article, chapter, or book as a final authority if a psychologist can imagine disagreeing with any of the findings or conclusions in some circumstances or characterizations.

Regarding essential sources of Rorschach research, several empirical studies exist that have summarized this information and have directly addressed previous criticisms of the measure. Before testifying, it would behoove psychologists to familiarize themselves with each of the original sources

that will be cited. Interrater reliability of the Rorschach has been found to be good to excellent based on accepted psychometric standards (Meyer, 1997a, 1997b; Meyer et al., 2002; Viglione & Hilsenroth, 2001). Test–retest reliability data for the Rorschach have been shown to be quite good and at least comparable with other psychological tests (Grønnerød, 2003; Viglione & Hilsenroth, 2001). Likewise, the Rorschach has demonstrated validity both broadly and in specific domains (Bornstein, 1996, 1999; Hiller, Rosenthal, Bornstein, Berry, & Brunell-Neuleib, 1999; Meyer & Archer, 2001; Meyer et al., 2001; Rosenthal, Hiller, Bornstein, Berry, & Brunell-Neuleib, 2001). This is especially true in relation to objectively determined outcome and ecologically valid behavioral criteria (Hiller et al., 1999; Meyer & Archer, 2001; Rosenthal et al., 2001).

Important Rorschach contributions to psychotherapy include indexes of termination, engagement, continuation, and prediction of success in psychotherapy (S. Ackerman, Hilsenroth, Clemence, Weatherill, & Fowler, 2000; Colson, Eyman, & Coyne, 1994; Hilsenroth, Handler, Toman, & Padawer, 1995; Meyer, 2000a; Meyer & Handler, 1997; Nygren, 2004). In addition, the Rorschach can be useful in selecting appropriate treatment modalities, in monitoring change and improvement over time on indexes of adjustment among both adolescents as well as adults receiving psychotherapy (Abraham, Lepisto, Lewis, Schultz, & Finkelberg, 1994; Bihlar & Carlsson, 2000, 2001; Blatt & Ford, 1994; Elfhag, Rossner, Lindgren, Andersson, & Carlsson, 2004; Exner & Andronikof-Sanglade, 1992; J. C. Fowler et al., 2004; Grønnerød, 2004; Stokes et al., 2003; Weiner & Exner, 1991). Validity has been demonstrated with specific clinical application to Axis II, Cluster B personality disorders (Blais, Hilsenroth, Castlebury, Fowler, & Baity, 2001; O'Connell, Cooper, Perry, & Hoke, 1989), aggressive ideation (Baity & Hilsenroth, 1999, 2002; Janson & Stattin, 2003), thought disorder (Hilsenroth, Fowler, & Padawer, 1998; Jørgensen, Andersen, & Dam, 2000; Perry & Braff, 1994, 1998; Perry, Geyer, & Braff, 1999; Perry, Minassian, Cadenhead, Sprock, & Braff, 2003; Viglione, 1999; Viglione & Hilsenroth, 2001), and suicidal activity (Exner & Wiley, 1977; C. Fowler, Piers, Hilsenroth, Holdwick, & Padawer, 2001).

It is important to note that in their summary and conclusion article of the *Psychological Assessment* special section on the Rorschach, Meyer and Archer (2001) provided extensive data comparing the Rorschach with the Wechsler Adult Intelligence Sacle (WAIS; Wechsler, 1997) and the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1943) in terms of a wide variety of validity approaches. Meyer and Archer's (2001) conclusion is explicit, bringing closure to the controversy about the Rorschach's clinical utility:

> There is no reason for the Rorschach to be singled out for particular criticism or specific praise. It produces reasonable validity, roughly on par with other commonly used tests. (pp. 491–492)

Meyer and Archer (2001) went on to state that validity is always conditional, a function of predictor and criterion, and that this limitation presents an ongoing challenge for all psychological assessment instruments. In sum, Rorschach reliability and validity is at least equivalent in comparison with other instruments in psychological and cognitive assessment.

A configurational, synthetic approach to Rorschach interpretation as described by Stricker and Gold (1999) and demonstrated by Weiner (1999) is fundamental to any clinical/actuarial method. This clinical/actuarial method with the Rorschach uses an ecologically valid and informed understanding of interactive probabilities to increase accuracy of in vivo decision making in applied psychology. Also, questions of why, when, and how one should rely on clinician-based inferences are much larger than the use of the Rorschach in forensic assessments. A recent review on these issues by Westen and Weinberger (2004; in press) integrates an emerging body of research on the psychometric quantification and adequacy of clinical observations and will be a valuable resource to any clinician preparing forensic testimony.

Exner (2001, 2002, 2003) provided extensive data for various nonclinical as well as clinical reference samples and is currently collecting an updated nonpatient sample. Exner (2001, 2002) discovered an error with a subsample of the original nonpatient records, and this error was corrected. Comparisons of the new nonpatient sample with the corrected sample of original protocols showed a great deal of similarity between the two data sets. Aspects of Exner's (1986, 1993) original nonpatient data have been questioned (Wood, Nezworski, Garb, & Lilienfeld, 2001), and some differences have been shown between the original data and current samples of adults and children (Hamel, Shaffer, & Erdberg, 2000; Shaffer, Erdberg & Haroian, 1999). Several factors have been identified that may account for the nature of the observed discrepancies, and a discussion of these criticisms in any context, forensic or otherwise (Meyer, 2001; Weiner, 2001), would be complex.

Chief among these factors is the nature and recruitment of the Exner (1986, 1993, 2002, 2003) nonpatient samples who were unpaid volunteers recruited through occupational, social, and community organizations as well as screened for any current and past psychiatric treatment. Only a very small percentage (17%) of those had received any psychiatric treatment, limited to eight or fewer sessions of academic, vocational, marital, or bereavement counseling. In addition, the new sample excludes any individual having a prolonged or significant use of psychotropic medication. Therefore, the inclusion rules for the Exner (1986, 1993, 2002, 2003) nonpatient samples are more restrictive (i.e., they require greater psychological health; see Kessler et al., 1994; U.S. Department of Health and Human Services, 1999) than those for "normative" samples for other personality assessment mea-

sures. In this sense, the Exner (1986, 1993, 2002, 2003) nonpatient samples might best approximate an asymptomatic reference group (Tingey, Lambert, Burlingame, & Hansen, 1996a, 1996b). Undoubtedly, additional normative samples using various recruitment strategies to include individuals in the community unscreened for past or current psychiatric treatment (psychotherapy or medication) or even given direct payment for their participation would be useful to clinicians. However, note that establishing a continuum of normative reference groups is certainly not standard practice for any instrument in the area of personality assessment, psychiatric interviewing, or cognitive functioning. Also, it is important for clinicians to examine the entire set of Exner's (2003) nonclinical and clinical reference samples when evaluating protocols (i.e., deviation from outpatient and inpatient as well as nonpatient data). As we have already discussed concerning the areas of reliability and validity (Grønnerød, 2003; Hiller et al., 1999; Meyer & Archer, 2001; Meyer et al., 2001; Rosenthal et al., 2001; Viglione & Hilsenroth, 2001), the extensive nonpatient and clinical reference groups provided to clinicians for the Rorschach Comprehensive System (CS; Exner, 2003) as well as the procedures used in obtaining them are best appreciated in context through comparison with other instruments in psychological and cognitive assessment.

Related to these general normative issues, the Rorschach has shown a lack of ethnic bias with regard to nonclinical and clinical groups of African Americans (Meyer, 2002; Presley, Smith, Hilsenroth, & Exner, 2001). In fact, any limited differences that were observed in a clinical sample actually favored minorities (i.e., less pathological) and worked against patients of Euro-American descent (i.e., more pathological; Meyer, 2002). In addition, principal components factor analysis of this clinical sample revealed no evidence of ethnic bias in the Rorschach's internal structure (Meyer, 2002). Related to ethnic diversity, Rorschach data has been provided in large-scale investigations of cultural diversity with international samples (Erdberg & Shaffer, 1999, 2001).

Although the evidence for the reliability and validity of Rorschach indexes are equivalent to other psychological assessment measures (Meyer & Archer, 2001), like any assessment measure, it also has limitations that a testifying expert should be aware of when questioned about. For instance, a recent review by Jørgensen et al. (2000) found very favorable evidence for the use of the Rorschach in the assessment of thought disorder, but this same review showed little support for the use of the Rorschach Depression Index (DEPI) in the categorical classification or discrimination of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed. [*DSM–IV*]; American Psychiatric Association, 1994) diagnosed depression. Therefore, extant data would caution clinicians against the use of the DEPI in the prediction of *DSM–IV* diagnoses of depression. Although the DEPI may not be effective in the assessment of *DSM–IV* diagnoses of depression, this does not mean all Rorschach variables are ineffective in such an endeavor. Recent research by Hartman, Wang, Berg, & Saether (2003) demonstrated that several Rorschach variables were useful in differentiating *DSM–IV* depressed, previously depressed, and never depressed individuals. In addition, logistic regression analyses indicated that several Rorschach variables (Wsum6%, X – %, X + %, C%, SumY%, MOR%, EBPer) significantly improved the prediction of *DSM–IV* major depression incrementally beyond information provided by the Beck Depression Inventory (BDI; Beck & Steer, 1987). Finally, the design of this study encompassed all procedural guidelines previously put forth by several Rorschach critics (Wood, Lilienfeld, Garb, & Nezworski, 2000).

In addition, Rorschach critics have repeatedly suggested that the Rorschach is being used in isolation to determine whether someone has been sexually abused (Garb, Wood, & Nezworski, 2000a, 2000b; Lilienfeld, Fowler, Lohr, 2003; Lilienfeld, Wood, & Garb, 2000; Lohr, Fowler, & Lilienfeld, 2002). We are not aware of any psychologists who would advocate making the determination of a sexual abuse history of an individual on the basis of the Rorschach alone. In fact, the argument against doing this is made quite explicitly by several authors who have investigated Rorschach differences between sexually abused and nonabused groups (Kamphuis, Kugeares, & Finn, 2000; Leavitt, 2000; Leifer, Shapiro, Martone, & Kassem, 1991) as well as in reviews of this issue that have attempted to correct misperceptions about Rorschach assessment as applied to victims of sexual abuse (Meyer & Archer, 2001; Weiner, Spielberger, & Abeles, 2002, 2003). It is important to note the difference between citing the empirical evidence showing that the Rorschach can provide useful information about abuse status at the group level and extrapolating these group differences to applied clinical decisions about abuse at the individual case level in the absence of other data. This same caution about drawing idiographic conclusions on the basis of nomothetic data applies to other areas of Rorschach interpretation as it would to any measure, sign, or index used for personality assessment.

Each of us has separately discussed elsewhere (Stricker & Gold, 1999; Viglione & Hilsenroth, 2001) the ability of the Rorschach to assess implicit personality characteristics (McClelland, Koestner, & Weinberger, 1989; Shedler, Mayman, & Manis,1993; Westen & Weinberger, 2004, in press). Although this is a strength of the method, it is also important to acknowledge the limitation of making inferences based on these implicit dynamics without also assessing the explicit personality structure of an individual. Because the Rorschach does provide idiographic information about the implicit qualities of an individual, such information is most responsibly utilized within an interpretive matrix containing information about explicit attitudes and overt behavior. It is by using the Rorschach in concert with methods evaluating conscious aspects of personality and observed behavior that a more complex and differentiated psychological assessment may be completed.

Controlled studies have demonstrated the effectiveness of several indexes across various assessment measures to detect significant differences between simulator and comparison groups (Lewis, Simcox, & Berry, 2002; Rogers, Sewell, Grandjean, & Vitacco, 2002; Rogers, Sewell, & Salekin, 1994). The Rorschach has demonstrated an ability to produce valid protocols from participants attempting to conceal psychological disturbance (i.e., emotional distress, self-critical ideation, and difficulties in interpersonal relationships) on self-report measures of personality (Ganellen, 1994; Ganellen, Wasyliw, Haywood, & Grossman, 1996; Grossman, Wasyliw, Benn, & Gyoerkoe, 2002). These findings may be of particular relevance to forensic and court-ordered assessment.

With specific regard to discussing the use of the Rorschach in forensic settings, there is no better summation of pertinent issues and refutation of likely criticisms that may be encountered than two recent articles by Ritzler, Erard, and Pettigrew (2002a, 2002b). Ritzler et al. (2002a, 2002b) specifically exposed several erroneous assertions by Grove and Bordon (1999) and Grove, Borden, Garb, and Lilienfeld (2002) that Rorschach-based testimony is inadmissible under legal standards. Ritzler et al. (2002a, 2002b) provided highly specific information concerning the actual expert testimony legal guidelines (i.e., *Daubert v. Merrell Dow Pharmaceuticals, Inc.,* 1993; *Federal Rules of Evidence* [*FRE*], 1992; *Frye v. United States,* 1923; *General Electric Co. v. Joiner,* 1997; *Kumho Tire Co., Ltd. v. Carmichael,* 1999).

Although the Rorschach has been a lightning rod for criticism, the forensic issues we discuss apply to any psychological assessment instrument. Several recent articles provide the same essential resources for an array of forensically relevant instruments including the Psychopathy Check List–Revised (Hare, 1991; Gacono, Loving, Evans, & Jumes, 2002; Loving 2002), Millon Clinical Multiaxial Inventory–III (MCMI–III; Millon, Davis, & Millon, 1997; McCann, 2002; Retzlaff, Stoner, & Kliensasser, 2002), MMPI–2 (Butcher, 1998; Butcher et al., 2001; Forbey & Ben-Porath, 2002; Otto, 2002), Personality Assessment Inventory (Edens, Cruise, & Buffington-Vollum, 2001; Morey, 1991; Morey & Quigley, 2002), and Bender Gestalt (Raphael, Golden, & Cassidy-Feltgen, 2002). In particular, information provided by these recent works are quite pertinent to the issues of test norms, reliability/validity, dissimulation/malingering, child custody evaluations, and interpretation of diagnostic characteristics.

CONTROVERSY DOES NOT EQUAL INADMISSIBILITY

First, controversy regarding a test does not constitute a specific criterion for establishing admissibility in court. In other words, courts clearly do not expect unanimous endorsement from a discipline or profession regarding a measure or viewpoint and realistically expect disagreement among experts (*Breda v. Wolf Camera,* 1998). One need only look to the issues of yearly mammograms for women over the age of 40, hormone replacement therapy, chronic fatigue syndrome, fibromyalgia, and the use of pharmacological agents (e.g., cholesterol medication for women vs. men, Ritalin, single or combined use of HIV medications, etc.) to find controversies in the field of medicine. The original principles on the admissibility of expert testimony in *Frye v. United States* (1923) provide that the basis of the expert opinion be based on procedures that have "general acceptance" in that discipline. However, a revision of *FRE* (1992) 702 (*FRE* 702) provided even more liberal parameters in federal and some state courts that include as admissible information that is "helpful" or provides "assistance" to the *trier of fact* (i.e., the judge or the jury) if it has acceptance even within a "specialized" community of professionals of a discipline (O'Conner & Krauss, 2001). Also, the U.S. Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993) decided that the helpfulness and relevance criteria found in *FRE* 702 superceded the general acceptance standard under *Frye v. United States* (1923), thus changing the primary emphasis from what other experts in the field believe to questions of adequate methodological rigor and "fit" with the issues in the particular case. General acceptance may still be considered as one of the "indicia" of helpfulness, but it is not ordinarily restrictive by itself.

It is important for the expert witness, guided by counsel, to be cognizant of the prevailing rules of evidence in the court where the psychologist is testifying. A number of states have not yet adopted the *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993) standard and continue to rely on *Frye v. United States* (1923) guidelines, and still others use rules of evidence that can differ in certain material respects from the *FRE* (1992). Although psychologists who practice in states where some variant of the *Frye v. United States* standard is utilized might have some of their test based testimony challenged on the basis of the controversy issue, the consulting attorney can argue that psychological tests such as the Rorschach or MMPI–2 are hardly novel and thus should be presumed to meet a general acceptance standard.

Some previous discussion regarding the requirements of psychological assessment measures has offered a radical new interpretation of legal guidelines (Garb, 1999; Grove & Barden, 1999; Grove et al., 2002; Wood, Nezworski, Stejskal, & McKinzey, 2001) without reference to relevant court cases or commentaries. It is not the case that any test used to help inform an expert opinion be established beyond any scientific controversy. It also is not the case, according to current legal standards, that acceptance by a majority of "academic, university scientists" is needed prior to the use of any test to form an expert opinion. Furthermore, a test need not demonstrate incremental validity with regard to specific forensic issues to be used in expert testimony. It is tempting for a critic to assert a wish as though it were a fact, but it re-

mains for the court, guided by the pertinent rules of evidence, to determine admissibility.

The legal standards regarding test admissibility in many jurisdictions state that experts must "show that they have followed the scientific method as it is practiced by (at least) a recognized minority of scientists in their field" (*Daubert v. Merrell Dow Pharmaceuticals, Inc.,* 1993, p. 1319). Contrary to the assertion (Grove et al., 2002) that academics in a university setting should define what is accepted in a discipline, in the *Kumho Tire Co., Ltd. v. Carmichael* (1999) decision, the Supreme Court stated, regarding issues of gate keeping, that "It is to make certain that an expert, whether basing testimony upon professional studies or personal experience, employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field" (p. 1176). It is especially important here to note the Supreme Court's focus on both the relevance of "personal experience" and "the practice" (e.g., applied clinical practice in the field of clinical psychology) of an expert, a point we turn to later in discussion of the clinical competencies of expert witnesses.

It is questionable to assume that any of these controversy issues actually apply to the Rorschach. Regarding the general acceptance of the Rorschach, repeated surveys of psychological test use over the past 40 years have shown a substantial, consistent, and sustained use of this instrument in academic training as well as in research and clinical settings. Surveys consistently indicate that more than 80% of graduate programs teach the Rorschach and that students regard this training as important in developing other clinical skills, in understanding their patients better, and as useful in their practicum/internship training (Archer & Newsom, 2000; Camara, Nathan, & Puente, 2000; Hilsenroth & Handler, 1995; Piotrowski & Zalewski, 1993; Watkins, Campbell, Nieberding, & Hallmark, 1995). Although some surveys have reported that the Rorschach takes more of the clinician's time than many other tests, they have also reported that it is one of the most frequently used (e.g., Ball, Archer, & Imhof, 1994; Camara et al., 2000). Also, 90% of clinical practitioners working in the field expressed a belief that clinical students should be competent in Rorschach assessment (Watkins et al., 1995). The two most comprehensive recent surveys of predoctoral internships (Clemence & Handler, 2001; Stedman, Hatch, & Schoenfeld, 2000), each including over 300 Association of Psychology Postdoctoral and Internship Centers sites (most of which were programs accredited by the American Psychological Association), revealed that internship training directors greatly value the Rorschach as well as integrated test batteries. Again, training directors reported a desire for incoming interns to have had courses or a good working knowledge of the Rorschach. Finally, the Rorschach is the second most frequently researched personality assessment instrument. This rate of annual research has been stable over a 20-year period, with a mean of more than 95 studies published per year (Butcher & Rouse, 1996).

Despite concentrated criticisms of the Rorschach (Garb, 1999; Garb, Wood, Nezworski, Grove, & Stejskal, 2001; Grove & Barden, 1999; Grove et al., 2002; Lilienfeld et al., 2003, 2000; Wood, & Lilienfeld, 1999; Wood et al., 2000; Wood, Lilienfeld, Nezworski, & Garb, 2001; Wood, Nezworski, Garb, et al., 2001; Wood, Nezworski, & Stejskal, 1996; Wood, Nezworski, Stejskal, Garven, & West, 1999; Wood, Nezworski, Stejskal, et al., 2001), conscientious attention to the available systematically gathered evidence does not support the existence of a widespread controversy concerning the Rorschach. Articles critical of the Rorschach written by the same cadre of authors do not represent evidence of wide-spread controversy in the field of personality assessment or clinical psychology in general. The empirical evidence does not support this skeptical position, yet these critics persist in their arguments. Given the empirical, clinical, training, and research data reported earlier, a majority (indeed, in all likelihood, a substantial majority) of the relevant professional community seems to view the Rorschach as a reliable, valid, and clinically useful measure to use in the training and practice of clinical psychology. Indeed, judgments about the utility of an assessment instrument should not be determined by a majority vote but rather by the weight of the data. The data concerning the reliability and validity of the Rorschach is substantial and at minimum, equivalent to other psychological assessment instruments used in forensic settings.

Similarly, recent surveys of professional patterns of psychological test usage have revealed the MMPI (and MMPI–2) and MCMI (Millon, 1983; and MCMI–II [Millon, 1987] and MCMI–III) to be the two most commonly used self-report personality tests in general clinical practice (Camara et al., 2000; Watkins et al., 1995) as well as in both civil (Boccaccini & Brodsky, 1999) and criminal (Borum & Grisso, 1995) forensic evaluations. In addition, surveys of child custody evaluations reveal a similar pattern of use (M. Ackerman & Ackerman, 1997; Quinnell & Bow, 2001).

EXPERTISE RESIDES IN THE PSYCHOLOGIST AND NOT IN THE ASSESSMENT INSTRUMENT

Expert testimony is qualified on the basis of the knowledge, skill, experience, training, and education that aids in forming that expert opinion. It is important to note that in almost all cases regarding testimony using psychological assessment, it is the professional's expert opinion that is figural and not the specific test methods (McCann, 1998, 2004; Meloy, Hansen, & Weiner, 1997). The courts understand that the expertise pertinent to a given case is that of the professional who uses a wide array of information in arriving at an expert opinion. When expert testimony is ruled inadmissible, this decision is usually based on the application or relevance of the testimony to the legal issue in question or on the misapplication

of information contributing to that testimony. It usually does not hinge on the specific assessment instruments used by the expert.

A simplistic understanding of the psychological assessment process (i.e., professional opinions are formed on the basis of simple, single-sign variables from one test used in isolation) leads to the position that the skill of the expert is not considered to be part of the expert testimony. In fact, some have argued that all psychological expert testimony should be based purely on statistical or actuarial prediction rather than derived from clinical judgment (Garb, 1998, 1999; Grove & Barden, 1999; Grove et al., 2002; Wood, Nezworski, Stejskal, et al., 2001). These authors have further suggested that the integration of data across multiple methods and sources will only contaminate or obscure accurate clinical judgments. A belief that every individual component score and interpretive principle must be validated in several different contexts for a measure finally to be acceptable to use in expert testimony is at odds with the multimethod approach to psychological assessment recommended by the American Psychological Association (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999; American Psychological Association Committee on Professional Practice Standards, 1994). Similarly, as noted by Ritzler et al. (2002a), an exclusive reliance on actuarial prediction would exclude most medical testimony, as the ordinary process of assessment in medicine also entails an experience-based integration of information from multiple sources. The Federal Judicial Center Reference Guide on Medical Testimony (Henefin, Kipen, & Poulter, 2000) clearly states that "medical diagnosis is not an exact science" (p. 465), "findings are not at all truly independent of one another" (p. 467), and "doctors combine probabilities of disease (prevalence) with their knowledge of the frequency of signs and symptoms in a given disease to progressively modify and ultimately arrive at their view of the likelihood of the disease under consideration" (p. 467). Finally, and perhaps most important to the discussion at hand, the *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993) standard does not require the expert to provide validity estimates of each isolated probability or to depart from standards used in everyday applied practice.

Regarding the specific use of the Rorschach in expert testimony, Weiner, Exner, and Sciara (1996) reviewed 93 survey responses from forensic psychologists who testified in more than 4,000 criminal cases, more than 3,000 custody cases, and 858 personal injury cases for a total of almost 8,000 cases. In only 6 of these cases did the respondents report that the integrity of the test was challenged, and in only one case was the psychologist's testimony ruled inadmissible. Similarly, Meloy et al. (1997) reviewed court of appeals citations from 1945 to 1995. Among the 194 cases addressing the Rorschach, for 89.5% "the admissibility and weight of Rorschach data were not questioned by either the appellant or the respondent and were important enough to be mentioned and discussed in the legal ruling by the court of appeal" (Meloy et al., 1997, p. 60). The Meloy et al. study found that it was the interpretation or findings derived from the test that were challenged. The use of the Rorschach per se was specifically attacked in only 2 cases. These data demonstrate that the Rorschach "has authority, or weight, in higher courts of appeal in the United States" (Meloy et al., 1997, p. 61).

In contrast to these prior articles, a recent survey by Lally (2003) might be used to suggest that the Rorschach is not widely used by forensic psychologists. In this survey, Lally reported the opinions of a small group of forensic psychologists (n = 40 to 53) working exclusively within criminal law settings (i.e., not custody evaluations, personal injury, etc.) regarding the acceptability of a variety of psychological tests in six areas of practice (mental state at the offense, risk for violence, risk for sexual violence, competency to stand trial, competency to waive *Miranda* rights, and malingering). The Rorschach was rated by a majority of respondents as unacceptable in five out of six of these areas. It is worthwhile to point out that the Rorschach does not provide data that are directly relevant to some of the specific forensic questions addressed in the Lally (2003) survey, and therefore, it is not surprising that it was not recommended for use. It was surprising, however, that these forensic psychologists rated the WAIS–III and MMPI–2 positively with regard to these forensic questions in the absence of any specific validity data. The Rorschach may not have been a good choice for these questions, but it is puzzling why the other tests were considered to be better.

In a systematic, comprehensive legal analysis of the admissibility of the Rorschach, McCann (1998) concluded that a focus on the structural summary from the CS "meets professional and legal standards for admissibility of psychometric evidence and expert testimony" (p. 140). Legal challenges to the use of the Rorschach in forensic assessment may well increase in the coming years given the very low rate reported in the empirical literature currently. Such legal challenges to the Rorschach will likely be a result of attorneys being encouraged to make these objections by some vocal critics in clinical psychology (Garb, 1998, 1999; Grove & Barden, 1999; Grove et al., 2002; Wood, Nezworski, Stejskal, et al., 2001; Ziskin, 1995). It will be valuable to review the outcome of these challenges several years from now, but to date, the Rorschach has easily met standards of admissibility in the legal process.

Challenges to the admissibility of expert testimony based on the MMPI/MMPI–2 have also been quite rare (Otto, 2002). As was the case with the Rorschach, when instances of admissibility of expert testimony based on the MMPI/MMPI–2 were challenged, it was the manner in which this testimony was used rather than the psychometric properties of the measure that was contested. A similar survey on admissibility concerning the MCMI–II/MCMI–III has yet to be conducted, but all available information sug-

gests that an equally high level of admissibility is enjoyed by this measure (McCann, 2002).

LEGAL DEBATE IS NOT "SCIENTIFIC" OR UNBIASED

The legal process consists of an adversarial presentation of arguments designed to arrive at a determination of the truth. Each attorney is expected to present an argument and only the evidence that supports that specific argument. The presentation of alternative arguments and evidence is the responsibility of opposing counsel. Accordingly, one should not expect an unbiased examination or discussion of the relevant assessment issues or the reporting of supportive evidence when it does exist (Barrett & Morris, 1993). There is no clearer example of this than the volumes of Ziskin and Faust (1988; see pp. xvii, xviii) and a later edition by Ziskin (1995; see pp. x, xi) alone criticizing psychiatric and psychological testimony. These authors explicitly discussed the adversarial nature of their volumes (i.e., being partisan and one-sided), which focus entirely on literature that negates the expertise of mental health professionals, actively excluding supportive literature of any kind:

> The exclusion of supportive literature is *not* to cause readers to believe it does not exist. As noted, it may or does exist. However, although perhaps of academic interest, such supportive evidence is viewed as largely irrelevant in a forensic context. (Ziskin, 1995, p. x).

This incontrovertible fact may actually aid the forensic psychologist being cross-examined by an attorney utilizing the Ziskin and Faust books. As Weiner (1996) so adroitly pointed out

> Psychologists need only observe that the work of doctors Ziskin and Faust is not based on a scientific approach to the data and does not involve a comprehensive or impartial review of the literature, which is an assertion that can be documented by reading the preface of their book. (p. 209)

In addition, the literature review in the fifth edition (Ziskin, 1995) volumes are out of date and contain conclusions that are inconsistent with the current scientific evidence. In a forensic setting, expert testimony based on the Ziskin volumes is extremely vulnerable to cross-examination by an attorney who has been familiarized with the current Rorschach literature.

Unfortunately, Rorschach critiques in the scientific literature seem to have adopted this (legal) adversarial model of argumentation in their template for scientific writing. These extreme positions, such as calling for a moratorium on the use of the test (Garb, 1999; Garb, Florio, & Grove, 1998, Garb et al., 2001; Grove & Barden, 1999; Grove et al., 2002; Wood, Nezworski, Stejskal, et al., 2001), suggesting that those who use the Rorschach may be practicing "pseudoscience" (Lilienfeld et al., 2003; Lohr et al., 2002) or are at risk for ethical complaints (Garb, 1999; Grove & Barden, 1999; Hunsley & Bailey, 1999, 2001; Lilienfeld et al., 2000; Wood, Nezworski, Stejskal, et al., 2001) and failing to acknowledge substantial empirical support for the measure despite an extensive, peer-reviewed, empirical literature (Garb, 1999; Garb et al., 2001; Grove & Barden, 1999; Grove et al., 2002; Lilienfeld et al., 2003; Lilienfeld et al., 2000; Lohr et al., 2002; Wood & Lilienfeld, 1999; Wood et al., 2000; Wood, Lilienfeld, et al., 2001; Wood, Nezworski, Garb, et al., 2001; Wood et al., 1996; Wood et al., 1999; Wood, Nezworski, Stejskal, et al., 2001) all seem more consistent with a legal model of argument than a scientific approach. In this same vein, more stringent standards seem to be used to evaluate studies that report positive Rorschach results than negative ones, and negative findings are often cited, whereas positive results in the same study are not (for numerous examples, see Viglione & Hilsenroth, 2001). In addition, Rorschach critics have not attended to direct feedback regarding mistakes in their analyses (see Meyer, 2000b), and when original data were provided to Rorschach critics (Dawes, 1999), questionable analyses were conducted, and important positive findings were not reported (see Perry, 2001, 2003). The issue is not that a group of scholars have reached conclusions different from ours but that subsequent arguments do not reflect the complexity of findings in the entire corpus of evidence. A one-sided presentation concerning any topic may have a place in politics or in the adversarial arena of the courtroom, but it is ill suited to scientific publication.

A pertinent example of an incomplete presentation of data in regard to forensic issues is the recently cited but unpublished manuscript by McKinzey and Campagna (2001, as cited in Brodsky & McKinzey, 2002; available at http://www.wpe.info). McKinzey and Campagna (2001) provided suggestions on how to rebut expert testimony using Rorschach findings in a provocatively titled manuscript "The Rorschach, Exner's Comprehensive System, Interscorer Agreement, and Death." Despite the provocative title, no standard reliabilty coefficients (e.g., intraclass correlation coefficient, kappa, etc.) for the protocol in question were presented, several of the citations in this study were used inappropriately, and several other studies were available at the time (i.e., before 2002) to McKinzey and colleagues that directly contradict their comments. Finally, contrary to how the information regarding this case is presented it was an issue of diagnosis, not the use of the Rorschach per se (i.e., the Rorschach was not ruled inadmissible), that was the focus of the appellate argument. Regardless of this significant limitation, Brodsky and McKinzey (2002) discussed aspects of this unpublished study including

> In another death penalty case (McKinzey & Campagna, 2001), an appellate argument was rebutted when it was

pointed out that a psychologist retained by defense counsel used the Rorschach Comprehensive System without taking into account its 81% false-positive rate (Mittman, 1983), discredited norms (Wood, Nezworski, Garb, & Lilienfeld, 2001), and questionable scoring system (Guarnaccia, Dill, Sabatino, & Southwick, 2001; Wood, Nezworski, & Stejskal, 1996).

It is important to note the lack of reference to literature demonstrating the diagnostic efficiency rates of various CS scores and indexes regarding assessment of psychosis (Jørgensen et al., 2000) or other clinical conditions (Fowler et al., 2001) in applied settings. In addition, the impact of standardized Rorschach administration, scoring, and interpretation in regard to life or death forensic contexts is of particular concern. We would agree that any forensic psychologist who obtained a poorly recorded Rorschach protocol is practicing far outside the bounds of what would be considered reasonable CS standards of administration. However, the Mittman (1983) reference to a "81% false-positive rate" is from an unpublished doctoral dissertation that concerned whether individuals could malinger schizophrenia and not the implied general validity of the "Rorschach Comprehensive System." The reference to the norms of the CS as "discredited" and the scoring system as "questionable" is the prerogative of Brodsky and McKinzey (2002) to state their opinion, but it is certainly not the prevailing sentiment in the field in light of the data presented earlier. Also, there were several articles available to them that directly responded to the work they cited as the basis of their claims (Hiller et al., 1999; Meyer, 1997a, 1997b, 2000a, 2000b, 2001, 2002; Meyer & Archer, 2001; Meyer et al., 2002; Rosenthal et al., 2001; Viglione & Hilsenroth, 2001; Weiner, 2001). It is surprising that such emphasis was placed on an unpublished manuscript (McKinzey & Campagna, 2001) that had been rejected from two different journals (*Journal of Clinical Psychology* and *Professional Psychology: Research and Practice*) with unanimously negative peer review letters from two separate action editors and six reviewers (see McKinzey, 2002). Unfortunately, this example illustrates how a trier of fact (i.e., the judge or the jury) might be misled by a one-sided presentation of data and reaffirms the need for forensic psychologists using any assessment measure to be aware of the original sources and the current research literature. Only then can the use of inaccurate and misleading information be exposed for the lack of comprehensive scholarship it represents.

EXAMINE THE TRAINING, EXPERIENCE, COMPETENCIES, AND CERTIFICATIONS OF ANY EXPERT

Qualification as an expert in legal proceedings often requires an extensive review of the expert's educational training, experience with assessment measures or clinical techniques pertinent to the testimony, publication record, organizational activity, and certifications/specialties (i.e., psychological assessment, forensic psychology). As noted by Gacono (2002a, 2002b; Gacono, Evans, & Viglione, 2002), this expertise is subject to cross-examination. Those adequately trained and having appropriate experience with any assessment measure are well prepared to provide such expert testimony. A witness's credibility may be weakened if there are deficits in training, experience, and competence in areas of purported expertise.

Certainly, all well-trained proponents of psychological assessment in forensic settings believe that the use of single variables, signs, or measures to determine any forensically relevant decision is ill advised. This is not the standard practice of clinicians who are adequately trained in psychological assessment. Nonetheless, these are exactly the straw man arguments made by several critics of the Rorschach (Garb, 1999; Garb et al., 2001; Grove & Barden, 1999; Grove et al., 2002; Lilienfeld et al., 2003; Lilienfeld et al., 2000; Lohr et al., 2002; Wood et al., 2000; Wood, Nezworski, Stejskal, et al., 2001) that have claimed widespread assessment abuses and need for radical designs to protect the integrity of the legal system.

The use of extreme criticisms regarding the reliability, validity, and subsequent admissibility of the Rorschach may well prove a liability to the side that uses such a strategy. This testimony may be easily impeached on cross-examination by a well-informed attorney (Gacono, 2002a, 2002b; Gacono, Evans, et al., 2002; Ritzler et al., 2002a, 2002b; Viglione & Hilsenroth, 2001). If these issues are explored in the courtroom, the limits on their credibility in presenting the implications of assessment findings should be apparent. Although there is no record of an ethical complaint being made against an adequately trained forensic psychologist for simply using the Rorschach in an assessment battery, the testimony of psychologists who have only academic credentials, are untrained in the use of a measure they desire to provide testimony on, are unlicensed, or are not actively engaged in the applied clinical practice in the field of clinical psychology may be challenged under the *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993) and *Kumho Tire Co., Ltd. v. Carmichael* (1999) standards (Ritzler et al., 2002a). Finally, forensic psychologists are encouraged to provide consultation to attorneys regarding the cross-examination of those psychologists who offer prejudicial testimony on assessment measures beyond their training, clinical experience, and competence.

CONCLUSIONS

In this article, we provided a summary of how the complex wealth of information contained in several recent articles can be used in a concerted manner to address challenges directed

against the use of psychological assessment measures in a forensic setting. Recently, the use of specific assessment instruments and psychological assessment in general has come under attack. These critiques portray the interpretation and application of test variables and measures in a concrete, simplistic, and isolated manner (e.g., one reflection response on the Rorschach indicates narcissistic personality disorder). Such oversimplified analogues to actual clinical hypothesis building, testing, and decision making are not clinically appropriate, representative, relevant, or meaningful (Westen & Weinberger, 2004, in press). Similarly, the empirical reviews conducted by many critics often lack the integrated and comprehensive understanding that would be found through an examination of the breadth of the extant empirical literature. Instead, the conclusions drawn appear to be polemical, negatively biased, insulated from contradicting research, and extreme with regard to suggestions for future assessment practices.

In contrast, several rebuttals to these extant criticisms have discussed the forensic utility of the Rorschach in the context of a responsible psychological assessment process as well as a thorough review of the essential empirical literature. Furthermore, each of these recent forensic articles have addressed the need to acknowledge the strengths and limitations of the measures reviewed, the need for direct information to be derived from a test in a manner appropriate to the relevant legal question, and the need for a multimethod approach to evaluating various dimensions of functioning. As has been previously discussed regarding reliability and validity (Grønnerød, 2003; Hiller et al., 1999; Meyer & Archer, 2001; Meyer et al., 2001; Rosenthal et al., 2001; Viglione & Hilsenroth, 2001) as well as the extensive reference groups provided to clinicians for the Rorschach CS variables (Exner, 2002, 2003; Exner et al., 2001), the credibility of Rorschach-based testimony should be apparent when points of comparison are made with other instruments in psychological and cognitive assessment (Meyer & Archer, 2001).

ACKNOWLEDGMENT

We acknowledge the helpful suggestions and contributions of three anonymous reviewers on an earlier version of this manuscript.

REFERENCES

Abraham, P., Lepisto, B., Lewis, M., Schultz, L., & Finkelberg, S. (1994). An outcome study: Changes in Rorschach variables of adolescents in residential treatment. *Journal of Personality Assessment, 62,* 505–514.

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

Ackerman, S., Hilsenroth, M., Clemence, A., Weatherill, R., & Fowler, C. (2000). The effects of social cognition and object representation on psychotherapy continuation. *Bulletin of the Menninger Clinic, 64,* 386–408.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychological Association Committee on Professional Practice Standards. (1994). Guidelines for child custody evaluations in divorce proceedings. *American Psychologist, 49,* 677–680.

Archer, R., & Newsom, C. (2000). Psychological test usage with adolescent clients: Survey update. *Assessment, 7,* 227–235.

Baity, M., & Hilsenroth, M. (1999). Rorschach aggression variables: A study of reliability and validity. *Journal of Personality Assessment, 72,* 93–110.

Baity, M., & Hilsenroth, M. (2002). Rorschach aggressive content variable: A study of criterion validity. *Journal of Personality Assessment, 78,* 275–287.

Ball, J., Archer, R., & Imhof, E. (1994). Time requirements of psychological testing: A survey of practitioners. *Journal of Personality Assessment, 63,* 239–249.

Beck, A., & Steer, R. (1987). *Beck Depression Inventory manual*. San Antonio, TX: Psychological Corporation.

Barrett, G., & Morris, S. (1993). The American Psychological Associations' amicus curiae brief in Price Waterhouse v. Hopkins: The values of science versus the values of the law. *Law and Human Behavior, 17,* 201–216.

Bihlar, B., & Carlsson, A. (2000). An exploratory study of agreement between therapist's goals and patient's problems revealed by the Rorschach. *Psychotherapy Research, 10,* 196–214.

Bihlar, B., & Carlsson, A. (2001). Planned and actual goals in psychodynamic psychotherapies: Do patient's personality characteristics relate to agreement? *Psychotherapy Research, 11,* 383–400.

Blais, M., Hilsenroth, M., Castlebury, F., Fowler, C., & Baity, M. (2001). Predicting *DSM–IV* Cluster B personality disorder criteria from MMPI–2 and Rorschach data: A test of incremental validity. *Journal of Personality Assessment, 76,* 150–168.

Blatt, S., & Ford, R. (1994). *Therapeutic change: An object relations perspective.* New York: Plenum.

Boccaccini, M., & Brodsky, S. (1999). Diagnostic test usage by forensic psychologist in emotional injury cases. *Professional Psychology: Research and Practice, 30,* 253–259.

Bornstein, R. F. (1996). Construct validity of the Rorschach Oral Dependency scale: 1967–1995. *Psychological Assessment, 8,* 200–205.

Bornstein, R. F. (1999). Criterion validity of objective and projective dependency tests: A meta-analytic assessment of behavioral prediction. *Psychological Assessment, 11,* 48–57. Borum, R., & Grisso, T. (1995). Psychological test use in criminal forensic evaluations. *Professional Psychology: Research and Practice, 26,* 465–473.

Breda v. Wolf Camera, 78 Fair Empl. Prac. Cas. 433 (S.D. Ga., 1998).

Brodsky, S., & McKinzey, R. (2002). The ethical confrontation of the unethical forensic colleague. *Professional Psychology: Research and Practice, 33,* 307–309.

Butcher, J. (1998). *Users guide for the Minnesota Report: Report for forensic settings.* Minneapolis, MN: NCS.

Butcher, J., Graham, J., Ben-Porath, Y., Tellegen, A., Dahlstrom, W., & Kaemmer, B. (2001). *MMPI–2: Manual for administration and scoring* (Rev. ed.). Minneapolis: University of Minnesota Press.

Butcher, J., & Rouse, S. (1996). Personality: Individual differences and clinical assessment. *Annual Review of Psychology, 47,* 87–111.

Camara, W., Nathan, J., & Puente, A. (2000). Psychological test usage: Implications in professional use. *Professional Psychology: Research and Practice, 31,* 141–154.

Clemence, A., & Handler, L. (2001). Psychological assessment on internship: A survey of training directors and their expectations for students. *Journal of Personality Assessment, 76,* 18–47.

Colson, D., Eyman, J., & Coyne, L. (1994). Rorschach correlates of treatment difficulty and of the therapeutic alliance in psychotherapy with fe-

male psychiatric hospital patients. *Bulletin of the Menninger Clinic, 58,* 383–388.

Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579, 113 S. Ct. 2786 (1993).

Dawes, R. (1999). Two methods for studying incremental validity of a Rorschach variable. *Psychological Assessment, 11,* 297–302.

Edens, J., Cruise, K., & Buffington-Vollum, J. (2001). Forensic and correctional applications of the personality assessment inventory. *Behavioral Sciences and the Law, 19,* 519–543.

Elfhag, K., Rossner, S., Lindgren, T., Andersson, I., & Carlsson, A. (2004). Rorschach personality predictors of weight loss with behavior modification in obesity treatment. *Journal of Personality Assessment, 83.*

Erdberg, P., & Shaffer, T. (1999, July). *International symposium on Rorschach nonpatient data: Findings from around the world.* Symposium presented at the 16th Congress of the International Rorschach Society, Amsterdam.

Erdberg, P., & Shaffer, T. (2001, March). *An international symposium on Rorschach nonpatient data: Worldwide findings.* Symposium presented at the annual convention of the Society for Personality Assessment, Philadelphia.

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

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

Exner, J. E., Jr. (2001). *A Rorschach workbook for the Comprehensive System* (5th ed.). Asheville, NC: Rorschach Workshops.

Exner, J. E., Jr. (2002). A new nonpatient sample for the rorschach comprehensive sample: A progress report. *Journal of Personality Assessment, 78,* 391–404.

Exner, J. E., Jr. (2003). *The Rorschach: A Comprehensive System: Vol. 1. Basic foundations* (4th ed.). New York: Wiley.

Exner, J. E., Jr., & Andronikof-Sanglade, A. (1992). Rorschach changes following brief and short-term therapy. *Journal of Personality Assessment, 59,* 59–71.

Exner, J. E., Jr., & Wiley, J. (1977). Some Rorschach data concerning suicide. *Journal of Personality Assessment, 41,* 339–348.

Federal rules of evidence. (1992). Boston: Little, Brown.

Forby, J., & Ben-Porath, Y. (2002). Use of the MMPI–2 in the treatment of offenders. *International Journal of Offender Therapy and Comparative Criminology, 46,* 308–319.

Fowler, J. C., Ackerman, S., Speanberg, S., Bailey, A. Blagys, M., & Conklin, A. (2004). Personality and sympton change in treatment-refractory inpatients: Evaluation of the phase model of change using Rorschach, TAT, and *DSM–IV* Axis V. *Journal of Personality Assessment, 83.*

Fowler, J. C., Piers, C., Hilsenroth, M., Holdwick, D., & Padawer, R. (2001). Assessing risk factors for various degrees of suicidal activity: The Rorschach suicide constellation (S–CON). *Journal of Personality Assessment, 76,* 333–351.

Frye v. United States, 292 F. 1013 (D.C. Cir. 1923).

Gacono, C. (2002a). Introduction to a special series: Forensic psychodiagnostic testing. *Journal of Forensic Psychology Practice, 2,* 1–10.

Gacono, C. (2002b). Why there is a need for the personality assessment of offenders. *International Journal of Offender Therapy and Comparative Criminology, 46,* 271–273.

Gacono, C., Evans, B., & Viglione, D. (2002). The Rorschach in forensic practice. *Journal of Forensic Psychology Practice, 2,* 33–54.

Gacono, C., Loving, J., Evans, B., & Jumes, M. (2002) The Psychopathy Checklist–Revised: PCL–R testimony and forensic practice. *Journal of Forensic Psychology Practice, 2,* 11–32.

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

Ganellen, R., Wasyliw, O., Haywood, T., & Grossman, L. (1996). Can psychosis be malingered on the Rorschach? An empirical study. *Journal of Personality Assessment, 66,* 65–80.

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

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

Garb, H., Florio, C., & Grove, W. (1998). The validity of the Rorschach and the Minnesota Multiphasic Personality Inventory: Results from meta-analyses. *Psychological Science, 9,* 402–404.

Garb, H., Wood, J., & Nezworski, M. (2000a). Projective techniques and the detection of child sexual abuse. *Child Abuse & Neglect, 24,* 437–438.

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

Garb, H., Wood, J., Nezworski, M., Grove, W., & Stejskal, W. (2001). Towards a resolution of the Rorschach controversy. *Psychological Assessment, 13,* 433–448.

General Electric Co. v. Joiner, 118 S. Ct. 512 (1997).

Grønnerød, C. (2003). Temporal stability in the Rorschach method: A meta-analytic review. *Journal of Personality Assessment, 80,* 272–293.

Grønnerød, C. (2004). Rorschach assessment of changes following psychotherapy: A meta-analytic review. *Journal of Personality Assessment, 83.*

Grossman, L., Wasyliw, O., Benn, A., & Gyoerkoe, K. (2002). Can sex offenders who minimize on the MMPI conceal psychopathology on the Rorschach. *Journal of Personality Assessment, 78,* 484–501.

Grove, W.,& Barden, R. (1999). Protecting the integrity of the legal system: The admissibility of testimony from mental health experts under *Daubert/Kumho* analysis. *Psychology, Public Policy, and the Law, 5,* 224–242.

Grove, W., Barden, R., Garb, H., & Lilienfeld, S. (2002). Reply to Ritzler, Erard, & Pettigrew. *Psychology, Public Policy, and the Law, 8,* 216–234.

Hamel, M., Shaffer, T., & Erdberg, P. (2000). A study of nonpatient preadolescent Rorschach protocols. *Journal of Personality Assessment, 75,* 280–294.

Hare, R. (1991). *The Hare Psychopathy Checklist–Revised.* Toronto, Ontario, Canada: Multi-Health Systems.

Hartman, E., Wang, C., Berg, M., & Saether, L. (2003). Depression and vulnerability as assessed by the Rorschach method. *Journal of Personality Assessment, 81,* 243–256.

Hathaway, S. R., & McKinley, J. C. (1943). *The Minnesota Multiphasic Personality Inventory.* Minneapolis: University of Minnesota Press.

Henefin, M. S., Kipen, H. M., & Poulter, S. R. (2000). Reference guide on medical testimony. In *Reference manual on scientific evidence* (pp. 439–484). Washington, DC: Federal Judicial Center.

Hiller, J., Rosenthal, R., Bornstein, R., Berry, D., & Brunell-Neuleib, S. (1999). A comparative meta-analysis of Rorschach and MMPI validity. *Psychological Assessment, 11,* 278–296.

Hilsenroth, M., Fowler, C., & Padawer, J. (1998). The Rorschach Schizophrenia Index (SCZI): An examination of reliability, validity, and diagnostic efficiency. *Journal of Personality Assessment, 70,* 514–534.

Hilsenroth, M., & Handler, L. (1995). A survey of graduate students' experiences, interests, and attitudes about learning the Rorschach. *Journal of Personality Assessment, 64,* 243–257.

Hilsenroth, M. J., Handler, L., Toman, K. M., & Padawer, J. R. (1995). Rorschach and MMPI–2 indices of early psychotherapy termination. *Journal of Consulting and Clinical Psychology, 63,* 956–965.

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

Hunsley, J., & Bailey, J. (2001). Wither the Rorschach? An analysis of the evidence. *Psychological Assessment, 13,* 472–485.

Janson, H., & Stattin, H. (2003). Prediction of adolescent and adult delinquency from childhood Rorschach ratings. *Journal of Personality Assessment, 81,* 51–63.

Jørgensen, K., Andersen, T., & Dam, H. (2000). The diagnostic efficiency of the Rorschach Depression Index and the Schizophrenia Index: A review. *Assessment, 7,* 259–280.

Kamphuis, J., Kugeares, S., & Finn, S. (2000). Rorschach correlates of sexual abuse: Trauma content and aggression indexes. *Journal of Personality Assessment, 75,* 212–224.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., et al. (1994). Lifetime and 12-month prevalence of *DSM–III–R* psychiatric disorders in the United States. *Archives of General Psychiatry, 51,* 8–19.

Kumho Tire Co., Ltd. v. Carmichael, 119 S. Ct. 1167 (1999).

Lally, S. (2003). What tests are acceptable for use in forensic evaluations? A survey of experts. *Professional Psychology, 34,* 491–498.

Leavitt, F. (2000). Surviving roots of trauma: Prevalence of silent signs of sex abuse in patients who recover memories of childhood. *Journal of Personality Assessment, 74,* 311–323

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

Lewis, J., Simcox, A., & Berry, D. (2002). Screening for feigned psychiatric symptoms in a forensic sample by using the MMPI–2 and the Structured Inventory of Malingered Symptomatology. *Psychological Assessment, 14,* 170–176.

Lilienfeld, S., Fowler, K., & Lohr, J. (2003). And the band played on: Science, pseudoscience, and the Rorschach Inkblot Method. *The Clinical Psychologist, 56,* 6–7.

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

Lohr, J., Fowler, K., & Lilienfeld, S. (2002). The dissemination and promotion of pseudoscience in clinical psychology: The challenge to legitimate clinical science. *The Clinical Psychologist, 55,* 4–10.

Loving, J. (2002). Treatment planning with the Psychopathy Checklist–Revised (PCL–R). *International Journal of Offender Therapy and Comparative Criminology, 46,* 281–293.

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. (2002). Guidelines for forensic application of the MCMI–III. *Journal of Forensic Psychology Practice, 2,* 55–70.

McCann, J. (2004). Projective assessment of personality in forensic settings. In M. Herson (Series Ed.) & M. Hilsenroth & D. Segal (Vol. Eds.), *Comprehensive handbook of psychological assessment: Vol. 2. Personality assessment* (pp. 562–572). New York: Wiley.

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review, 96,* 690–702.

McKinzey, R. (2002, April 30). Prior peer review of "Rorschach interscorer agreement." *WebPsychEmpiricist.* Retrieved July 3, 2002, from http://www.woe.info/papers_table.html

McKinzey, R., & Campagna, V. (2002, April 27). The Rorschach, Exner's Comprehensive System, interscorer agreement, and death. *WebPsychEmpiricist*. Retrieved July 3, 2002, from http://www.wpe.info/papers_table.html

Meloy, J., Hansen, T., & Weiner, I. (1997). Authority of the Rorschach: Legal citations during the past 50 years. *Journal of Personality Assessment, 69,* 53–62.

Meyer, G. (1997a). Assessing reliability: Critical correlations for a critical examination of the Rorschach Comprehensive System. *Psychological Assessment, 9,* 480–489.

Meyer, G. (1997b). Thinking clearly about reliability: More critical correlations regarding the Rorschach Comprehensive System. *Psychological Assessment, 9,* 495–498.

Meyer, G. (2000a). Incremental validity of the Rorschach Prognostic Rating scale over the MMPI Ego Strength Scale and IQ. *Journal of Personality Assessment, 74,* 356–370.

Meyer, G. (2000b). On the science of Rorschach research. *Journal of Personality Assessment, 75,* 46–81.

Meyer, G. (2001). Evidence to correct misperceptions about Rorschach norms. *Clinical Psychology: Science & Practice, 8,* 389–396.

Meyer, G. (2002). Exploring possible ethnic differences and bias in the Rorschach Comprehensive System. *Journal of Personality Assessment, 78,* 104–129.

Meyer, G., & Archer, R. (2001). The hard science of Rorschach research: What do we know and where do we go. *Psychological Assessment, 13,* 486–502.

Meyer, G., Finn, S., Eyde, L., Kay, G., Moreland, K., Dies, R., et al. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist, 56,* 128–165.

Meyer, G., & Handler, L. (1997). The ability of the Rorschach to predict subsequent outcome: A meta-analysis of the Rorschach Prognostic Rating Scale. *Journal of Personality Assessment, 69,* 1–38.

Meyer, G., Hilsenroth, M., Baxter, D., Exner, J., Fowler, C., Piers, C., et al. (2002). An examination of interrater reliability for scoring the Rorschach Comprehensive System in eight data sets. *Journal of Personality Assessment, 78,* 219–274.

Millon, T. (1983). *MCMI manual.* Minneapolis, MN: National Computer Systems.

Millon, T. (1987). *Manual for the MCMI–II* (2nd ed.). Minneapolis, MN: National Computer Systems.

Millon, T., Davis, R., & Millon, C. (1997). *MCMI–III manual* (2nd ed.). Minneapolis, MN: National Computer Systems.

Mittman, B. (1983). Judges' ability to diagnose schizophrenia on the Rorschach: Effect of malingering. *Dissertation Abstracts International 44*, 4B. (Digital Dissertations AAT 8315540)

Morey, L. (1991). *The Personality Assessment Inventory.* Odessa, FL: Psychological Assessment Resources.

Morey, L., & Quigley, B. (2002). The use of the Personality Assessment Inventory (PAI) in assessing offenders. *International Journal of Offender Therapy and Comparative Criminology, 46,* 333–349.

Nygren, M. (2004). Rorschach Comprehensive System variables in relation to assessing dynamic capacity and ego strength for psychodynamic psychotherapy. *Journal of Personality Assessment, 83.*

O'Connell, M., Cooper, S., Perry, C., & Hoke, L. (1989). The relationship between thought disorder and psychotic symptoms in borderline personality disorder. *Journal of Nervous and Mental Disease, 177,* 273–278.

O'Conner, M., & Krauss, D. (2001). Legal update: New developments in Rule 702. *APLS News, 21,* 1–4, 18.

Otto, R. (2002). Use of the MMPI–2 in forensic settings. *Journal of Forensic Psychology Practice, 2,* 71–92.

Perry, W. (2001). Incremental validity of the Ego Impairment Index: A re-examination of Dawes (1999). *Psychological Assessment, 13,* 403–407.

Perry, W. (2003). Let's call the whole thing off: A response to Dawes (2001). *Psychological Assessment, 15,* 582–585.

Perry, W., & Braff, D. (1994). Information-processing deficits and thought disorder in schizophrenia. *American Journal of Psychiatry, 151,* 363–367.

Perry, W. & Braff, D. (1998). A multimethod approach to assessing perseverations in schizophrenia patients. *Schizophrenia Research, 33,* 69–77.

Perry, W., Geyer, M. A., & Braff, D. L. (1999). Sensorimotor gating and thought disturbance measured in close temporal proximity in schizophrenic patients. *Archives of General Psychiatry, 56,* 277–281.

Perry, W., Minassian, A., Cadenhead, K., Sprock, J., & Braff, D. (2003). The use of the Ego Impairment Index across the schizophrenia spectrum. *Journal of Personality Assessment, 80,* 50–57.

Piotrowski, C., & Zalewski, C. (1993). Training in psychodiagnostic testing in APA-approved PsyD and PhD clinical psychology programs. *Journal of Personality Assessment, 61,* 394–405.

Presley, G., Smith, C., Hilsenroth, M., & Exner, J. (2001). Rorschach validity with African Americans. *Journal of Personality Assessment, 77,* 491–507.

Quinnell, F., & Bow, J. (2001). Psychological tests used in child custody evaluations. *Behavioral Sciences and the Law, 19,* 491–501.

Raphael, A., Golden, C., & Cassidy-Feltgen, S. (2002). The Bender-Gestalt Test (BGT) in forensic assessment. *Journal of Forensic Psychology Practice, 2,* 93–106.

Retzlaff, P., Stoner, J., & Kliensasser, D. (2002). The use of the MCMI–III in screening and triage of offenders. *International Journal of Offender Therapy and Comparative Criminology, 46,* 319–333.

Ritzler, B., Erard, R., & Pettigrew, G. (2002a). A final reply to Grove and Barden: The relevance of the Rorschach Comprehensive System for expert testimony. *Psychology, Public Policy, and the Law, 8,* 235–246.

Ritzler, B., Erard, R., & Pettigrew, G. (2002b). Protecting the integrity of Rorschach expert witnesses: A reply to Grove and Barden (1999) re: The admissibility of testimony under Daubert/Kumho analysis. *Psychology, Public Policy, and the Law, 8,* 201–215.

Rogers, R., Sewell, R., Grandjean, N., & Vitacco, M. (2002). The detection of feigned mental disorders on specific competency measures. *Psychological Assessment, 14,* 177–183.

Rogers, R., Sewell, R., & Salekin, K. (1994). A meta-analysis of malingering on the MMPI–2. *Assessment, 1,* 227–237.

Rosenthal, R., Hiller, J., Bornstein, R., Berry, D., & Brunell-Neuleib, S. (2001). Meta-analytic methods, the Rorschach, and the MMPI. *Psychological Assessment, 13,* 449–451.

Schaffer, T., Erdberg, P., & Haroian, J. (1999). Current nonpatient data for the Rorschach, WAIS–R, and MMPI–2. *Journal of Personality Assessment, 73,* 305–316.

Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist, 48,* 1117–1131.

Stedman, J., Hatch, J., & Schoenfeld, L. (2000). Preinternship preparation in psychological testing and psychotherapy: What internship directors say they expect. *Professional Psychology: Research and Practice, 31,* 321–326.

Stokes, J., Pogge, D., Powell-Lunder, J., Ward, A., Bilginer, L., & DeLuca, V. (2003). The Rorschach Ego Impairment Index: Prediction of treatment outcome in a child psychiatric population. *Journal of Personality Assessment, 81,* 11–19.

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

Tingey, R., Lambert, M. J., Burlingame, G. M., & Hansen, N. (1996a). Assessing clinical significance: Proposed extensions to method. *Psychotherapy Research, 6,* 109–123.

Tingey, R., Lambert, M. J., Burlingame, G. M., & Hansen, N. (1996b). Clinically significant change: Practical indicators for evaluating psychotherapy outcome. *Psychotherapy Research, 6,* 144–153.

U.S. Department of Health and Human Services. (1999) *Mental health: A report of the Surgeon General.* Rockville, MD: Author.

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

Viglione, D., & Hilsenroth, M. (2001). The Rorschach: Facts, fictions, and future. *Psychological Assessment, 13,*452–471.

Watkins, C., Campbell, V., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychological Research and Practice, 26,* 54–60.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: Psychological Corporation.

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

Weiner, I. (1999). What the Rorschach can do for you: Incremental validity in clinical applications. *Assessment, 6,* 327–338.

Weiner, I. (2001). Considerations in collecting Rorschach reference data. *Journal of Personality Assessment, 77,* 122–127.

Weiner, I., & Exner, J. (1991). Rorschach changes in long-term and short-term psychotherapy. *Journal of Personality Assessment, 56,* 453–465.

Weiner, I., Exner, J., & Sciara, A. (1996). Is the Rorschach welcome in the courtroom? *Journal of Personality Assessment, 67,* 422–424.

Weiner, I., Spielberger, C., & Abeles, N. (2002). Scientific psychology and the Rorschach Inkblot Method. *The Clinical Psychologist, 55,* 7–12.

Weiner, I., Spielberger, C., & Abeles, N. (2003). Once more around the park: Correcting misinformation about Rorschach assessment. *The Clinical Psychologist, 56,* 8–9.

Westen, D., & Weinberger, J. (2004). When clinical description becomes statistical prediction. *American Psychologist, 59.*

Westen, D., & Weinberger, J. (in press). In praise of clinical judgment: Meehl's forgotten legacy. *Journal of Clinical Psychology.*

Wood, J., & Lilienfeld, S. (1999). The Rorschach Inkblot Test: A case of overstatement? *Assessment, 6,* 341–349.

Wood, J., Lilienfeld, S., Garb, H., & Nezworski, M. (2000). The Rorschach Test in clinical diagnosis: A critical review with a backward look at Garfield (1947). *Journal of Clinical Psychology, 56,* 395–430.

Wood, J., Lilienfeld, S., Nezworski, M., & Garb, H. (2001). Coming to grips with negative evidence for the comprehensive system for the Rorschach: A comment on Gacono, Loving, and Bodholt; Ganellen and Bornstein. *Journal of Personality Assessment, 77,* 48–70.

Wood, J., Nezworski, M., Garb, H., & Lilienfeld, S. (2001). The misperception of psychopathology: Problems with the norms of the comprehensive system for the Rorschach. *Clinical Psychology: Science and Practice, 8,* 350–373.

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

Wood, J., Nezworski, M., Stejskal, W., Garven, S., & West, S. (1999). Methodological issues in evaluating Rorschach validity: A comment on Burns and Viglione (1996), Weiner (1996), and Ganellen (1996). *Assessment, 6,* 115–129.

Wood, J., Nezworski, M., Stejskal, W., & McKinzey, R. (2001). Problems of the Comprehensive System for the Rorschach in forensic settings: Recent developments. *Journal of Forensic Psychology Practice, 1,* 89–103.

Ziskin, J. (1995). *Coping with psychiatric and psychological testimony* (5th ed.). Los Angeles: Law & Psychology.

Ziskin, J., & Faust, D. (1988). *Coping with psychiatric and psychological testimony* (4th ed.). Los Angeles: Law & Psychology.

Mark J. Hilsenroth
220 Weinberg Building
The Derner Institute of Advanced Psychological Studies
Adelphi University
Garden City, NY 11530
E-mail: hilsenro@adelphi.edu

Received August 15, 2002
Revised February 28, 2004