What the Rorschach Can Contribute to Child Custody and Parenting Time Evaluations

Robert E. Erard

SUMMARY. Personality assessment can indirectly help custody evaluators answer important questions about parental competence and the fit between parents’ psychological resources and children’s needs. It is preferable to use testing to check hypotheses derived from the case information rather than the reverse. The Rorschach makes it possible to assess implicit motives, coping capacities, and need states, thus complementing and supplementing self-attributed personal characteristics from the MMPI-2 and other self-report data. It also offers idiographic infor-
mation for individualizing assessments. Concerns about admissibility and charges that the Rorschach “overpathologizes” litigants are addressed, and suggestions are offered for making optimal use of the Rorschach in child custody work. [Article copies available for a fee from The Haworth Document Delivery Service: 1-800-HAWORTH. E-mail address: <docdelivery@haworthpress.com> Website: <http://www.HaworthPress.com> © 2005 by The Haworth Press, Inc. All rights reserved.]

KEYWORDS. Rorschach, child custody evaluation, forensic, norms, MMPI

Nearly all child custody and parenting time evaluations involve some kind of psychological testing (Ackerman & Ackerman, 1997; Keilin & Bloom, 1986; Quinnell & Bow, 2001). Survey data indicate that custody evaluators spend an average of about five hours in this activity (Ackerman & Ackerman, 1997; Bow & Quinnell, 2001; Keilin & Bloom, 1986), comprising about 25-30% of the total time they spend conducting the evaluations (not including report writing).

One reason that testing is such a common feature of these evaluations is, of course, that the people who write orders for them expect it to be done. Indeed, it is by no means unusual for evaluators to receive court orders that require “testing” without even mentioning the critical context in which it is embedded (i.e., the need for evaluation and recommendations for custody or parenting time). In recent surveys (described in Ackerman, Ackerman, Steffen, & Kelley-Poulos, 2004), between 73.9% and 85.5% of attorneys and judges expected psychological testing of parents or children to be administered by custody evaluators. In another survey (Bow & Quinnell, 2004), the mean ratings by attorneys and judges of psychological testing on a 1-5 scale of importance were 4.0. It is certainly plausible, nay likely, that consumers of custody and parenting time evaluations tend to overestimate both the validity and the relevance of psychological tests for this kind of legal decision-making. Their aura of scientific objectivity and precision and simply the fact that psychological tests are one source of data that cannot be duplicated in ordinary courtroom procedures probably play a large part in determining their popularity among legal professionals.

While there are many good scientific, ethical, and legal reasons not to base child custody recommendations entirely or even primarily on the results of psychological tests (APA COPPS, 1994; Bricklin, 1999;
Brodzinsky, 1993; Gould, 2004; Grisso, 1990; Heilbrun, 1992; Melton, Petrilka, Poythress, & Slobogin, 1997; Roseby, 1995; Stahl, 1999; Weithorn & Grisso, 1987), most psychologists concur with legal professionals in considering at least some personality testing of the parents to be a critical component of the generally accepted multisource, multimethod approach to data gathering (APA COPPS, 1994; Matarazzo, 1990; Weithorn & Grisso, 1987). Of course, psychologists know that there are few complex questions that can be confidently answered on the basis of psychological testing, unless they are integrated into a broader psychological assessment.

THE USE OF PERSONALITY TESTING IN CHILD CUSTODY EVALUATIONS

Personality testing of parents in child custody evaluations typically includes the MMPI-2 (Butcher, Graham, Ben-Porath, Tellegen, and Dahlstrom, 2001): 91.5% in the Ackerman and Ackerman (1997) survey and 94% in Quinnell and Bow (2001); the Rorschach, which is used somewhat less often: 47.8% in Ackerman and Ackerman (1997) and 44% in Quinnell and Bow (2001); and the MCMI series (Millon, Millon, and Davis, 1994: 34.3% in Ackerman and Ackerman (1997) and 52% in Quinnell and Bow (2001).

Whereas test selection by custody evaluators has often been surveyed, the professional literature offers somewhat less guidance about how custody evaluators ought to go about deciding which tests to use. Some authors (Brodzinsky, 1993; Underwager & Wakefield, 1993) have offered many instructive examples of how testing has been misused, but discussions of what specific contributions particular tests can make to the evaluation process are more challenging to find. Part of the problem is that, as Ellis (2000) has observed: “The majority of tests that are well validated and standardized have no direct bearing on parenting issues. The newer tests that have been developed to measure parenting skills or parent-child attachment . . . have not yet been rigorously validated” (p. 145).

Otto and Collins (1995) reviewed the MMPI/MMPI-2 literature and found no particular profile that identified parents as abusers or that was associated with abuse. Still more to the point, they were unable to identify any single scale or pattern of scales that was directly related to effective or ineffective parenting. A similar survey of the literature on the
Rorschach or the MCMI-III would scarcely be expected to yield more promising predictors of abuse or parental effectiveness.

Even though they are only indirectly informative about the specific psycholegal questions at issue in a child custody case, there is still good reason to use general personality measures. As noted in the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999), “...many tests measure constructs that are generally relevant to the legal issues even though norms specific to the judicial...context may not be available” (p. 129). Putting the case more concretely, Graham (2000) has observed that if a person has been described in his or her personality testing as “impulsive, unstable, unpredictable, and aggressive and as having very poor judgment, this would certainly be relevant to how this person might be expected to function in the parental role” (p. 376).

The particular relevance of personality testing in child custody evaluations is largely attributable to the social contexts in which divorce cases are selected for child custody evaluations. Because custody evaluations are costly, time-consuming, stressful, and emotionally draining, they tend to be ordered only in exceptional family law cases. These are primarily cases in which serious questions have been raised concerning parental fitness or the capacity of the parents to cooperate.

While parental unfitness may be the result of cognitive or physical limitations or a lack of relevant experience, most allegations of unfitness (or relative unfitness) that are referred for psychological evaluation involve problems such as emotional instability, self-centeredness, antisocial attitudes and behaviors, deficient empathy, aggressiveness, poor impulse control, and irrational beliefs or thought processes—in short, clinical personality issues. Questions concerning the capacity of parents to cooperate also tend to extend beyond the usual “irreconcilable differences” or “destruction of the objects of matrimony” that lead to divorce; the usual concern is about whether some basic characteristics of the parents’ individual makeup—their personalities—are very likely to prevent such cooperation over time. Somewhat less frequently, evaluations are also ordered because of concerns about the fit between one or both parents’ characteristic motives, values, and cognitive and emotional resources (again—their personalities) and the specific needs and characteristics (often, but not always, the personalities) of their children.

Thus, there is really no need to bemoan the fact that there are currently no psychological tests that are known to predict validly what cus-
tody or parenting time arrangement will be in the best interest of a particular child. Even though this is most often the ultimate issue that the expert has been asked to address, it is simply not the sort of question one should expect a personality test to answer. What constitutes a child’s “best interest” is, in the final analysis, not a psychological question, but a question of value to be determined over time in a socio-legal decision-making process guided by community standards along with statutes and case law.

What personality testing in particular can answer are often highly relevant intermediate questions along the way to an opinion about a child’s best interest: questions about parents’ capacity to regulate their affects and tolerate frustration, how sensitive they may be in understanding someone else’s feelings or a contrasting perspective, how realistic their expectations of themselves and others seem to be, whether they allow themselves to experience a full range of emotions and to respond to such feelings in others, what kinds of characteristic defenses and coping strategies they use when encountering intense emotional demands, how well they adapt to breaks in routine and changes in circumstances, and how much of their attention is focused on their own needs and feelings vs. those of other people. Questions of similar importance can be asked of personality testing in children: What cognitive, emotional, and social resources does the child have for coping with the particular strains of this divorce? How much stress or insecurity is the child experiencing and where is it showing up in her functioning? Is the child experiencing a developmental lag or succumbing temporarily to situational stress (Roseby, 1995)? How independent and adaptable is this child likely to be in a setting of rapidly shifting routines and environments? How mature is this child’s cognitive and social judgment in appraising his home environment? How sensitive does this child seem to be to family issues involving aggression, loss, sexuality, or unpredictability in the home environment?

In addition, personality tests bring to custody evaluations a kind of scientific anchoring for the subjectivity of clinical judgment. Features of personality tests such as standardized methods and conditions for gathering information, the availability of standardized norms and comparison groups, and the use of scientifically validated rules for preliminary interpretation offer the evaluator (and by extension, the trier of fact) a means of checking and quantifying subjective impressions and applying nomothetic decision rules to the data.

Unfortunately, it turns out that the degree of scientific certainty is usually greatest for relatively trivial and obvious questions and usually
weakest for especially complex, subtle, and interesting questions. As Edwin Schneidman (2004) puts it, “In all scientific endeavors, there is always a trade-off between precision and relevance, and we should be very wary of losing relevance for specious accuracy (even if it brings kudos and money)” (p. 133). While two qualified psychologists looking at the same personality test results will generally agree in broad outline about their meaning, the specific application of those results to a particular context of legal decision-making involves a significant degree of subjective judgment.

**TESTS AS SOURCES OF “HYPOTHESES”**

It is currently fashionable for custody evaluators (perhaps in imitation of the several standard computerized test report narratives that do so) to describe their findings from personality testing as “hypotheses,” rather than as facts or even opinions. It is often not clear what they expect the readers of these reports to do with such hypotheses. Ordinarily, one thinks of hypotheses as propositions that are going to be tested through some future research, but surely to leave it up to the finder of fact to devise some means of testing these “hypotheses” is an abdication of the expert’s own responsibility to offer reasonable conclusions as to whether or not these “hypotheses” apply in the instant case.

This is not to deny that it is sometimes a reasonable procedure in conducting an evaluation to use psychological tests as a source of hypotheses to be confirmed or refuted by other sources of data, such as personal and family histories, clinical observation, interview data, documentary records, collateral sources, and the like. Once such hypotheses have been confirmed to some reasonable degree of certainty, however, they ought to be offered as opinions. Alternatively, if they are not strong enough to stand on their own and cannot be confirmed by other data, there is no good reason to include them in the evaluation report—“hypotheses” or otherwise. Mere speculation has no useful role to play in the courtroom.

Irrelevant test findings are equally out of place. As Bricklin (1999) has observed, evaluators should consider whether information about the conceptual target that a test aims to measure, even if highly accurate, would make a difference regarding some decision: “If the answer is it would not, there is no reason to analyze a score’s value” (p. 125). What is most important is to avoid what Roseby (1995) has called a “pathology hunt,” where minor discrepancies from test norms are used to hold
custody litigants to higher standards of mental health than are generally applied to the non-litigating population.

However common it may be to derive clinical hypotheses from the test data and then look for factual confirmation, it is nevertheless probably not the best procedure in child custody evaluations. If the evaluator’s goal is (as it should be) to conduct a legally relevant evaluation, it stands to reason that the most important hypotheses to consider ought to originate in the factual and conceptual problems posed by the case itself (see Krishnamurthy, 2004).

In such an approach, one begins with the matter to be decided (e.g., whether a change of custody is in the best interest of the children, how parenting time should be divided, whether supervised parenting time is indicated) and then proceeds to the factually relevant questions. Such questions will often involve an evaluation of psychologically relevant factual allegations made by the parents that form an essential part of the legal basis for their pleadings. These may include claims about psychological characteristics of the other parent that might be associated with abuse or neglect of the children (such as low frustration tolerance, antisocial attitudes and behaviors, substance abuse) or that may interfere with the other parent’s capacity to co-parent (e.g., obsessive or delusional preoccupations about a former spouse that are being shared with the children).

Personality testing of children may also usefully address questions raised about the children’s particular needs and vulnerabilities (e.g., whether a child’s need for routine or current attachment status are such that a proposed parenting time arrangement would be too disruptive; whether a child’s emotional brittleness at home and poor school performance bespeak a maladaptive reaction to an untenable family situation; whether a child’s rigid adherence to one parent’s view of reality over the other’s is associated with anxiety about losing that parent’s love; whether a child is emotionally and socially resilient enough to adapt well to a relocation to a distant country, far away from one of her parents). From such questions, meaningful and relevant hypotheses can be derived about what sorts of data one should expect to see (and not see) in the personality testing and elsewhere.

Thus, the usual flow of inquiry is reversed. Rather than looking for discrepancies from population norms in the testing and generating a variety of possible concerns—shotgun style—and finally trying to shoehorn them into the legal issues posed by the case, one begins with the most important concerns flowing from the legal issues and generates differ-
ential hypotheses to be answered by the available data, including, of course, the testing.

Such hypothesis generation and testing are iterative. Each tentative observation and conclusion leads to further possibilities and new conceptualizations, which are subjected to further testing and confirmation. In this process, factual data from the family history and direct clinical observation come first, whereas testing and its associated theoretical constructs are used to refine one’s understanding of what the factual findings mean. As Fischer (1994) has proposed, “Our primary data are particular life events. Test scores, theoretical constructs, and research literature are all derived, secondary data through which the assessor modifies and refines his or her perception of life events—both those that are reported by the client and those that occur during the assessment” (p. 46).

In contrast to the test-centered approach, one that begins with hypotheses generated by the actual and implicit referral questions has two main advantages. First, it is more likely to emphasize the most relevant data, rather than what may be most psychometrically salient, thereby demonstrating in a transparent fashion what the findings mean in the context of the legal proceedings. By framing test results in terms of actual issues in the case, it becomes easier to construct a “narrative bridge” (Brodsky, 2004, p. 21) between the technical findings and the actual experience of the litigants. Second, from a Bayesian standpoint, such an approach maximizes inferential validity by keeping the conditional probabilities in the proper sequence. For example, instead of asking, given an elevated Scale 8 on the MMPI-2 or a high X-% on the Rorschach, “What is the likelihood that this father uses poor judgment in making decisions about his child?” the evaluator asks, “Given observational and historical data suggesting that father makes poor decisions in his parenting, is there test evidence that would tend to confirm that this is part of a more general problem with perceiving events realistically and conventionally?”

**WON’T THE MMPI-2 SUFFICE?**

Many forensic psychologists limit their personality testing to self-report measures such as the MMPI-2. There are several advantages in choosing this instrument. It is almost universally accepted by clinicians, researchers, and courts as a reliable and valid clinical personality test. It takes little clinician time to administer and is easy to score with automated methods. While the MMPI-2 lends itself to conceptually sophis-
icated and nuanced interpretations by masters of the instrument, even journeyman test users can routinely obtain reasonably valid results using cookbook methods.

In contrast, effective use of the Rorschach Comprehensive System (Exner, 1993) requires establishing an effective “testing relationship” with the subject and paying close attention to subtle indications to begin and end response inquiry. Administration and scoring alone often require two or more hours of clinician time. The Rorschach requires mastery of highly complex scoring rules and of interpretive routines that require simultaneous analysis of a wide array of variables and scoring configurations. Novices using cookbook methods may easily commit substantive interpretive errors. Maintenance of expertise in using the Rorschach for forensic purposes also generally requires many hours of devotion to developments in the literature concerning emerging scoring and interpretive methods, empirical foundations and range of application, and academic controversies. Why would anyone think it is worth the trouble?

Part of the answer has to do with the limitations of self-report tests, such as the MMPI-2. Particularly in high stakes situations, such as child custody evaluations, self-serving misrepresentations (Lanyon, 2001) in litigants’ statements about themselves are more or less the norm. Custody litigants taking self-report tests tend to ask themselves: “Would a parent who deserves to be awarded custody answer this question true or false?” Thus, it is hardly surprising that Bagby, Nicholson, Buis, Radovanovic, and Fidler (1999) found that “. . . Underreporting of symptoms in child custody litigation evaluations represents a real and significant challenge to psychologists using the MMPI-2 in their assessments of such litigants” (p. 28). Similar concerns have been voiced by Ollendick and Otto (1984), Siegel (1996), and Bathurst, Gottfried, and Gottfried (1997). While validity indicators can detect many naïve attempts at impression management, within-normal-limits profiles can be misleading in custody cases even in the absence of defensiveness on the validity scales (Graham, 2000). The Rorschach can often detect serious pathology that the MMPI-2 will miss (Ganellen, 1994, 1996).

When MMPI-2 results are treated as “hypotheses” to be confirmed by other “convergent data,” a still more subtle snare awaits the unwary evaluator. Most of the other types of data available to custody evaluators for comparison to MMPI-2 data, such as answers to questionnaires and in structured interviews, descriptions of personal history and family life, self-assessments of parenting skills and values, and answers on tests such as the Personality Assessment Inventory (Morey, 1996) and
the MCMI-III, are treated as multiple sources, but they are all mono-
method approaches in the sense that they are primarily based on self-re-
port. Even data gathered from the other party, the children, and collat-
eral sources shares several important method characteristics with verbal
self-report data; all of it requires verbal formulation and conceptualiza-
tion from the speaker (method similarity) and all of it involves at least
some dim awareness of how what is said may help or hurt the cause of a
given party (demand characteristics). When “convergent validity” is
established almost exclusively through self-report information, a great
deal of noise is introduced into the data in the form of illusory correlations
based on method similarities (Meyer, 1996, 1997, 1999) and demand
characteristics.

The Rorschach often leads to interpretations at variance with those
derived from MMPI-2 data (Archer & Krishnamurthy, 1993). This has
sometimes led critics (e.g., Wood, Nezworski, & Stejskal, 1996) to ar-
gue that the Rorschach must be inaccurate or not very useful. However,
its lack of routine correlation with self-report data is in fact one of its
most important virtues in forensic work. Because the Rorschach does
not rely on verbal self-report data and instead involves direct observa-
tion of performance under standardized conditions by the evaluator, it
can be an effective means of introducing method variance into the
evaluation process.

The two tests often complement each other, and the use of both to-
gether can provide incremental validity, particularly with regard to find-
ings about personality functioning (Archer & Krishnamurthy, 1993;
has written an excellent guide to the joint use of the MMPI-2 and the
Rorschach, and Weiner (1999) offers a striking case example of how the
tests may profitably be used together in a child custody evaluation.

When heteromethod convergence is achieved between self-report
measures and Rorschach findings, a high degree of confidence in such
findings is often justified. When the results are divergent or even con-
tradictory, this too can be clinically useful, because it may prompt the
evaluator to look more carefully at all available data in an effort to re-
solve the apparent contradiction. In some cases, the resolution may be
found in adjusting one’s interpretation of the meaning of a scale or
variable (e.g., an elevated MMPI-2 Scale 8 with a normal Percep-
tual-Thinking Index on the Rorschach may reflect profound alienation
and social insecurity rather than psychosis). In other cases, it may be
that the subject was less defensive or more engaged on one test or the
other (Ganelen, 1994; Meyer, 1999; Meyer, Riethmiller, Brooks,
Benoit, & Handler, 2000; Shapiro, Leifer, Marton, & Kassem, 1990). In still others, what might appear to be a personality trait on one test may be revealed as more likely the product of situational influences on the other (e.g., “paranoid” attitudes, suggested by MMPI-2 Scale 6 = 67, in the Rorschach context of T = 1, PTI = 0, COP = 3, AG = 0, and high M + %, are more likely to be interpreted as situational distrust).

Robert Bornstein (2001, 2002) has offered an intriguing analysis of the differential contributions of self-report and performance (projective) methods. According to Bornstein (2001), the Rorschach Comprehensive System is designed primarily to assess three things: implicit motives (need states that the subject cannot describe directly), cognitive/perceptual style, and aspects of the individual’s coping style such as stress tolerance and coping resources.

The concept of process dissociation has to do with the first of these: implicit motives. Bornstein cites McClelland, Koestner, and Weinberger (1989), who proposed that implicit motives predict an individual’s spontaneous behavior in less structured situations. Tests that tap into implicit motives like the Rorschach:

provide a more direct readout of motivational and emotional experiences than do self-reports that are filtered through analytic thought and various concepts of self and others. . . . Implicit motives are more often built on early, prelinguistic affective experiences whereas self-attributed motives are more often built on explicit teaching by parents and others as to what values or goals it is important for a child to pursue. (McClelland et al., 1989, pp. 698-699)

Bornstein (2002) reviews research contrasting measures of self-attributed motives with performance-based measures of implicit motives in several areas, including achievement motivation, intimacy, power, and dependency. Borrowing the concept of “process dissociation” from the learning and memory literature (where it has classically been applied to the problem of implicit vs. explicit memory), he explains how implicit and self-attributed motivations can be determined to represent different aspects of the same construct by demonstrating converging behavioral predictions, modest positive intercorrelations, and differential effects of moderating variables. He offers evidence that by administering self-report (e.g., MMPI) and performance (e.g., Rorschach) measures of a particular motive or need state to the same individual and studying their discontinuities, clinicians can obtain a more complete picture of that person’s underlying and expressed strivings.
A further advantage of including the Rorschach along with the MMPI-2 is the wealth of idiographic information it can offer, above and beyond the formal scoring relationships. While caution should be exercised in using idiographic data to make predictive statements, findings from a disciplined Rorschach content and sequence analysis, when integrated with biographical and nomothetic data, can help individualize the assessment (Fischer, 1994) by enriching the evaluator’s understanding of the dynamics, habitual defenses, coping capacities, resiliency, preoccupations, and other aspects of the inner life of the person being tested (Bricklin, 1999; Peebles-Kleiger, 2002; Weiner, 2000).

**IS THE RORSCHACH STILL WELCOME IN THE COURTROOM?**

In child custody and parenting time cases, the Rorschach seems to be competing with the MCMI-III for second place among the most commonly used personality tests, where it used to be firmly ensconced just below the MMPI-2 in popularity (see Quinnell & Bow, 2001). Despite the Rorschach’s unique advantages, some evaluators are reluctant to use it because they are afraid that Rorschach testimony may be found inadmissible in the courtroom. In support of such fears, they refer to a spate of articles over the past half dozen years attacking the use of the Rorschach in forensic practice.

From the standpoint of the legal system, if the Rorschach were in serious danger of being expelled from the courtroom, one might expect to see some trend in this direction in the nation’s appellate courts. Meloy, Hansen, and Weiner (1997), using a Lexis search, found 494 appellate citations of the Rorschach between 1945 and 1995, with specific discussions of the use of the test in 194 cases. In nearly 90% of these cases, “the admissibility and weight of the 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” (p. 60). While some courts questioned the specific relevance of findings from the testing, the validity of the method itself was attacked in only two cases.

Lower court decisions are not usually published, but Weiner, Exner, and Sciara (1996) surveyed 93 forensic psychologists about nearly 8,000 of their cases from across the U.S., 3,052 of which were child custody cases. The use of the Rorschach was challenged in only six cases, and in only one case was Rorschach-related testimony ruled inadmissi-
ble. This survey is currently being updated to see whether anything has changed in 2004.

Some Rorschach critics (Garb, 1999; Grove & Barden, 1999; Grove, Barden, Garb, & Lilienfeld, 2002; Wood, Nezworski, Garb, & Lilienfeld, 2001) are pinning their hopes that the Rorschach will eventually be banned from the courtroom on the current requirement in all federal courts and many state courts that judges make their own determinations about the scientific reliability of the methods on which experts base their testimony, following the U.S. Supreme Court cases, Daubert v. Merrell Dow Pharmaceuticals (1993) and Kumho Tire v. Carmichael (1999). While it may be true that psychologists who use the Rorschach Comprehensive System and other psychological tests will increasingly be called upon to justify their methods using the criteria set forth under Daubert (Medoff, 2003), this need not be very difficult if they have become conversant with the literature broadly defending the psychometric credentials of the Comprehensive System (e.g., Acklin, 1999; Meyer, 2000; Meyer & Archer, 2001; Rosenthal, Hiller, Bornstein, Berry, & Brunell-Neuleib, 2001; Viglione, 1999; Viglione & Hilsenroth, 2001) and defending the specific applicability of the Rorschach in forensic settings (Gacono, Evans, and Viglione, 2002; Hamel, Gallagher, & Soares, 2001; Hilsenroth & Stricker, 2004; McCann, 1998; Ritzler, Erard, & Pettigrew, 2002a, 2002b; Weiner, 1999, 2002), along with the work of the various critics cited earlier.

**DOES THE RORSCHACH OVERPATHOLOGIZE?**

One of the most virulent criticisms of the use of the Rorschach Comprehensive System in forensic settings is the claim that reliance on the standard Exner norms (Exner, 2001) leads to overpathologizing interpretations. In a number of U.S. and international samples of normal or “control” subjects, the average non-patient may appear to be psychologically unhealthy when compared to the Exner normative sample (see, for example, Shaffer, Erdberg, & Haroian, 1999; Wood, Nezworski, Garb, & Lilienfeld, 2001). Based on comparison to the Exner norms, more of these non-patients than expected could reasonably be considered to be narrow-minded and emotionally avoidant, to have mildly impaired reality testing, to be narcissistic, and to have limited capacity for warmth or emotional closeness with others. Some authors (Hunsley & DiGiulio, 2001; Wood, Nezworski, Garb et al., 2001;
Wood, Nezworski, Stejskal, & McKinsey, 2001; but see Meyer, 2001; Weiner, 2001a) have therefore questioned whether the Rorschach should even be used until newer, more representative norms have been established by research.

Several explanations have been offered for the failure of many local, so-called normal samples to pass Exnerian muster, including differences in recruiting and screening procedures (Exner’s are rather strict and may be biased toward particularly ego-involved, self-censorious, high-functioning volunteers–see Weiner, 2001a, 2001b), statistical drift (aggregated data from local convenience samples tend to differ systematically from national published norms–see Meyer, 2001), differences in the demographic characteristics between samples, increased stringency in the Comprehensive System scoring rules over time, local differences in administration (especially in inquiry) and in scoring, and a tendency for the population to become increasingly neurotic over time (see generally, Meyer, Viglione, Erdberg, Exner, & Shaffer, 2004).

Thus, the differences between the Exner norms and other samples are probably the product of several factors, some reflecting real population differences and others artifactual. The fact remains, however, that most contemporary studies using the Rorschach with non-patients obtain mean scores that fall somewhere between Exner’s non-patient norms and Exner’s outpatient samples (Meyer et al., 2004).

In response to questions about the continuing relevance of the Exner norms, John Exner is about to publish a new normative sample. Preliminary results (Exner, 2002) seem to fall somewhere in between the original norms and contemporary local samples on several key variables. Most variables commonly associated with serious symptoms or personality disorders (e.g., PTI, DEPI, CDI, HV1) remain reassuringly rare in Exner’s new normative data.

Four considerations seem to be particularly important in considering Rorschach norms and the assessment of psychopathology in child custody work. The first is: How representative are the Exner norms of scores likely to be obtained from nonpatients presenting for assessment? It is probably useful to think of the Exner norms as representing “healthy normals” as opposed to “people plucked at random off the street.” A sample consisting of the latter can be expected to contain a fair amount of psychopathology. For example, in the MMPI-2 restandardization sample, which is closer to the “person-in-the-street” population, 39% of the “normal” men and 43% of the “normal” women have clinically significant elevations on at least one MMPI-2 clinical scale (Shaffer & Erdberg, 2004). When compared to international Ror-
schach samples, the Exner normative sample looks about 0.4 standard deviations healthier than the rest (Meyer et al., 2004). In other words, the Exner norms seem “off” by about 4 T-score points, about the same as the difference between the MMPI and the MMPI-2 norms. It is likely that the new Exner norms will carve away at these differences.

Second, it is crucial to bear in mind that responsible Rorschach users do not draw sweeping conclusions about psychopathology based on simple difference scores between particular Rorschach variables and their non-patient normative references. While such discrepancies, when sizable, are usually taken into consideration when interpreting a Rorschach, no diagnostic or interpretive statement is justifiable by such findings alone. As Exner (1993) has taken pains to explain:

There is no simple checklist of Rorschach signs that automatically can be translated as representative of aspects of personality or behavior. . . . That supposition is pure nonsense, and the unwitting interpreter who is bound to that procedure would be better to avoid the use of the Rorschach in assessment work. The test is far too complex for a simplistic sign approach, and the validity of the interpretation will depend largely on the extent to which the interpreter achieves the full measure of data integration in forming the description of the subject. (p. 332)

Third, custody evaluators should try to become familiar with typical scoring patterns in the population of particular interest to them—viz., custody litigants. The effort to describe custody norms is already well underway for the MMPI-2 (Bagby et al., 1999; Bathurst et al., 1997) and the MCMI-III (Halon, 2001; Lampel, 1999; McCann, Flens, Campagna, Collman, Lazzarro, & Connor, 2001). Several preliminary studies have described typical Rorschach findings in custody cases (Bonieskie, 2000; Hoppe & Kenney, 1994; Lee, 1996; Singer, 2001), but so far none has made its way into the peer-reviewed literature. Because the purpose (and thus, meaning to the subject) of the testing is different from the ordinary clinical or research situation and because custody litigants may be facing unusual situational stressors, such comparisons may offer insight into what is usual and unusual for the situation of the custody litigant. Still, specialized norms only take us so far. As Graham (2000) has warned, the standard norms should remain the principal basis for comparison in all personality testing until extra-test correlates of scores for specialized groups are developed.
Fourth, psychologists who use the method described earlier, of looking first to the historical and observational data in the context of questions relevant to the legal determination and then using the testing for checking hypotheses and enriching one’s understanding of the real-world issues, will tend to minimize any tendencies to misperceive or exaggerate pathology based on discrepancies from norms.

**HOW TO USE (AND HOW NOT TO USE) THE RORSCHACH IN CHILD CUSTODY CASES**

Rorschach protocols used for forensic purposes should always be administered and scored according to the rules of the Comprehensive System in order to meet acceptable criteria for psychometric reliability and validity. Users of the Rorschach in child custody matters should be reasonably conversant with both the classic and contemporary scientific literature on the test and regularly attend workshops or symposia on its forensic applications.

They should strive to maintain a high level of expertise in administering and scoring the instrument, since standardized administration and scoring are the foundation of valid use. They ought to regularly check their administration and scoring competence (see Standard 5.9, *Standards for Educational and Psychological Testing*, AERA et al., 1999) by reviewing the instructions and examples from the most current *Workbook* (Exner, 2001) and testing themselves against the exercises in the Appendix. Donald Viglione’s (2002) recent guide to difficult scoring problems is also an invaluable resource.

There is good evidence that well-trained Rorschach scorers can achieve high levels of interrater reliability (Meyer, Hilsenroth, Baxter, Exner, Fowler, Piers et al., 2002). Even so, because even very experienced Rorschach users make occasional scoring or calculation errors, the use of a scoring program such as the RIAP 5 (Psychological Assessment Resources, Inc., Odessa, FL) or ROR-SCAN (Phil Caracena, Laredo, TX) is often helpful in preventing careless mistakes.

The forensic use of the associated interpretive programs is more problematic in two respects. First, as caveats in the narrative boilerplate of both commercial programs emphasize, an adequate interpretation for a given individual cannot be based on the formal scoring alone. Second, forensic experts who use such interpretive programs as part of the basis for their opinions may be ethically and legally required to produce them.
upon request, even though they may easily be mischaracterized by cross examiners.

Just as some MMPIs are more interpretable and useful than others, the same is true for Rorschachs. The most clinically reliable Rorschachs are those in which the test taker has understood the task and demonstrated adequate engagement with it, which can be determined by a combination of clinical observation and statistical features (particularly $R$ and Lambda). Rorschachs in which test takers have demonstrated an abnormal “set,” offered too few or too many responses, or failed to use a reasonable range of common determinants must, like MMPIs with questionable validity scale patterns, be interpreted with considerable caution, if at all.

Interpretation can be validly approached in a variety of ways, but for most users, systematic application of the interpretive procedures of the Rorschach Workshops Primer (Exner, 2000), along with careful attention to interpretable test behavior, projective content, and sequence analysis, following Irving Weiner’s (2000, 2003) carefully delineated principles, is the optimal procedure.

Of particular importance in forensic work are Weiner’s (2000) cautions (inter alia): (1) that examiners must first obtain a meaningful understanding of the underlying personality processes in a particular case before drawing clinical inferences about the test taker, (2) that they must try to attend to adaptive capacities as well as maladaptive tendencies, (3) that they should frame interpretive statements at appropriate levels of confidence, and (4) that they need to take the test taker’s cultural context into consideration.

To these might be added the warning that in custody work, the Rorschach should not be used for fishing expeditions into every conceivable area of psychopathology. Rather, as discussed earlier, examiners should begin with an understanding of the unresolved issues from the “story” of the case and come to the Rorschach with meaningful questions that it may be well suited to answer.

In contrast to the finite number of appropriate uses of the Rorschach, there are, of course, an infinite number of ways not to use it. Those of particular concern (because they are not as rare as they should be in child custody work) include: (1) relying on unscored or incorrectly scored Rorschachs or otherwise departing from the fundamental procedures of the Comprehensive System; (2) interpreting Rorschachs exclusively according to computerized interpretive programs; (3) attempting to determine the actual quality of relationships between parents and children based on Rorschach responses; (4) basing opinions about the best interests of the children exclusively or even primarily on
Rorschach results; (5) basing psychodiagnostic conclusions about litigants or family members (particularly children—see Hamel, Shaffer, & Erdberg, 2000) on the Rorschach alone; (6) determining whether children are victims of sexual abuse or whether parents are abusers primarily from the Rorschach (Kuehnle, 1996); and (7) presuming a fixed, universal significance for particular Rorschach plates or response contents.

Another common, but less obvious, problem is the assumption of a fixed, universal meaning for certain Rorschach determinants and indexes. For example, a Reflection response (Fr/rF) does not always signify a Narcissistic Personality Disorder; an absence of Texture (T/TF/FT) does not always point to a cool, unaffectionate parent; three or more white space (S) responses do not always signify an angry, defiant person; a reasonable number of Human Movement (M) responses does not guarantee a capacity for accurate empathy; not all test takers who produce Lambda greater than 1.0 have Avoidant Personality Disorder; many people who are clinically depressed do not elevate DEPI; even some people with an Adjusted D greater than or equal to 0 lack an adequate capacity to manage stress; not everyone who produces a Vista (V/VF/FV) response loathes herself; and many, if not most, physically abusive parents do not produce Aggressive Movement (AG) responses. The meaning of Rorschach scores and indexes in a particular case must be determined contextually—both by comparing them to and integrating them with convergent and contrasting data elsewhere in the Structural Summary and by considering their particular meaning in the life history and recent experiences of the test taker.

The Rorschach is a sensitive instrument that can allow child custody evaluators privileged access to important data about test takers’ coping capacities, cognitive and perceptual style, internal representations of self and others, and implicit motives and need states, using a methodology that does not rely on the test taker’s self-attributed personal characteristics. When it is applied within its own proper scope by a well-trained, experienced examiner, it is a formidable resource for coloring in the lines of a thorough personality assessment.

REFERENCES


Ackerman, M. J., Ackerman, M. C., Steffen, L. J., & Kelley-Poulos, S. (2004). Psychologists’ practices compared to the expectations of family law judges and attorneys in child custody cases. *Journal of Child Custody, 1*(1), 41-60.


