A FINAL REPLY TO GROVE AND BARDEN
The Relevance of the Rorschach Comprehensive System
For Expert Testimony

Barry Ritzler
Long Island University

Robert Erard
Psychological Institutes of Michigan, P.C.

Gary Pettigrew
Eastern Louisiana Mental Health System

In reply to Grove et al. (2002), the authors attempt to limit their focus on the question of admissibility of the Rorschach Comprehensive System for expert testimony under the guidelines of the U.S. Supreme Court Daubert/Kumho/Joiner decisions. The article refutes the argument that a “raging controversy” exists as evidence that the Rorschach is not accepted in the field of psychology. The authors again argue that Grove et al. have misconstrued the intent of Daubert/Kumho and misidentify nonclinician academics as the appropriate evaluators of the admissibility of the Rorschach. The authors add to their previous argument (2002) that the Rorschach has sufficient reliability, validity, and error rates to be admissible under Daubert and conclude by countering the Grove et al. argument that the Journal of Personality Assessment is not an adequate forum for peer review of the Rorschach.

With the addition of two members of a familiar group of Rorschach critics, the arguments by Grove et al. (2002) no longer center on the Daubert/Joiner/Kumho Supreme Court decisions but shift to what the authors consider to be wrong with the Comprehensive System in general. A few new criticisms are introduced, but much of the rejoinder repeats arguments previously published elsewhere by members of the same team of critics. While the original Daubert/Joiner/Kumho issues are given short shrift, the authors add a new consideration for expert testimony involving the Rorschach; i.e., their designation of the university professor as arbiter of the admissibility of expert testimony—a consideration absent from the Supreme Court’s decisions on expert evidence.

Grove et al. (2002) appear to have lost track of the context of their original article and our reply. We are not attempting to mount a comprehensive defense against all criticisms to which the Rorschach has been subjected in the scientific and professional literature. Instead, we limit ourselves to the central issues bearing on the admissibility of expert opinions based on Rorschach assessment. Most of the scattered criticisms Grove et al. now raise have already been definitively answered in the special series in Psychological Assessment (1999, 2001) and the Journal of Personality Assessment (2001).

Throughout their rejoinder, Grove et al. refer to their several critiques and those of some of their colleagues as half of a “raging controversy” (p. 229), an
“impassioned defense” (p. 216), an “intense scientific controversy” (e.g., p. 216), and “heated exchanges” (p. 217). While such rhetoric might simply be attributed to a flamboyant writing style, these characterizations cannot be dismissed because the authors use them as evidence that the Rorschach Comprehensive System has a highly questionable status for use in expert testimony. They conclude: “...the obviously strident ongoing controversy is evidence [our italics] of the lack of general acceptance by the relevant scientific community” (p. 229). From our perspective, this “intense scientific controversy” consists of the aforementioned critiques by a small team of authors with occasional replies by different respondents; i.e., mainly researchers who have conducted their own studies of the Comprehensive System. In our opinion, these circumstances do not define a “controversy” that is “raging” across the profession. This is the first of several examples of extreme statements by Grove et al. who offer no empirical support for such claims.

In terms of the present discussion of the application of the Daubert opinion, the recognition that the existence of some controversy surrounding a theory or technique, far from being an indictment of it, is an indication that it belongs to the living, breathing, growing part of an applied science, rather than to the dead wood of obsolete or discredited ideas. It is precisely through the dialog between proponents of a theory or technique and its critics that advancement in science takes place. As we argue further below, the opening of the courtroom to healthy scientific debate (and a concomitant effort to move away from a conception of expertise as stewardship of a storehouse of stagnant knowledge) was actually a central objective of the Supreme Court in developing the Daubert standard.

Our fundamental disagreement with Grove et al. concerns the interpretation of the Daubert/Joiner/Kumho guidelines by the Supreme Court justices. Grove et al. submit that Daubert and its progeny require that any technique used to form the basis of an expert’s opinion be established beyond scientific controversy. They argue that Daubert (and Frye v. United States before it) requires that the theory or technique forming the basis of an expert’s testimony be accepted by a majority of professors in the associated academic discipline. According to Grove et al., “It seems warranted to conclude that a majority (indeed, in all likelihood, a substantial majority) of the relevant scientific community...does not view the RCS as a reliable system with either broad-based zero-order validity or forensically useful incremental validity” (p. 217). No citations are offered to substantiate this viewpoint. Did Grove et al. survey the “relevant scientific community”? Their citations of the literature critical of the Rorschach consist of the same 8 or 10 authors (including themselves), listed in various permutations. Again, where is the empirical evidence “warranting” their disparaging conclusion?

Grove et al. operationally define “the relevant scientific community” for purposes of determining “general acceptance” as “…the predominantly academic, university scientists who conduct and review...personality assessment [research]” (p. 217). With respect to testimony about the Rorschach in particular, one must take this to mean university scientists who conduct and review Rorschach research. Because Grove et al. stipulate that these legitimate evaluators should be those who “conduct and review” [our italics] Rorschach studies, they have probably eliminated themselves from this community. A search of the PsycINFO database over the past 25 years does not reveal that they have
conducted studies of the Rorschach; their scholarship has been limited to criticism.

The relevant professional and scientific community for applying Daubert/Kumho guidelines to Rorschach-based opinion testimony is not and should never be the one advocated by Grove et al. Daubert was intended as a liberalization of the Frye test. The court took the position that the former “general acceptance” rule treated science as an established collection of knowledge, rather than a work in progress, whereas it held that Federal Rule of Evidence (FRE) 702 has a “more liberal thrust,” encouraging greater admissibility of expert testimony. The Daubert court acknowledged that the Federal Rules of Evidence “allow district courts to admit a somewhat broader range of scientific testimony that would be admissible under Frye” (Daubert v. Merrell Dow Pharmaceuticals, Inc., 1993, p. 2795).

Thus, it was under the Frye standard, not Daubert, that the academic members of a scientific community served as the primary gatekeepers, because testimony based on the discretion of members of that community was generally accepted under Frye (Krauss & Sales, 1999). Yet even under the more rigid Frye standard, close scientific scrutiny generally was limited to novel theories or techniques, not to whatever techniques of those practicing members of a professional community that had received the unanimous imprimatur of the academic branch of their field.

In effect, Daubert intentionally invited scientific controversy by firmly establishing that the judge, not scientists, should decide whether the proffered evidence meets reasonable criteria for admissibility in a particular case. As we noted in our original article (Ritzler, Erard, & Pettigrew, 2002), Ninth Circuit Judge Alex Kozinski, in his decision on remand in Daubert, found that admissible testimony by experts requires that they “can show that they have followed the scientific method as it is practiced by (at least) a recognized minority of scientists in their field” (p. 1319).

While neither Daubert v. Merrell Dow Pharmaceuticals, Inc. (1993) nor General Electric Co. v. Joiner (1997) addressed nonscientific expert testimony, in Kumho Tire Co. v. Carmichael (1999), the Supreme Court made it clear that it intended a similar analysis to take place in such cases. Some authors have taken Kumho to mean that testimony from applied disciplines (e.g., clinical psychology, medicine, engineering, accounting), where such testimony is generally based on some combination of scientific knowledge and experience-based practical judgment, experts must meet the most stringent scientific criteria on matters such as the known error rate of their applied judgment in order for their opinions to be admissible. However, the Kumho decision emphasized that “the test of reliability is ‘flexible,’ and Daubert’s list of specific factors neither necessarily nor exclusively applies to all experts or in every case” (Kumho Tire Co. v. Carmichael, 1999, pp. 1167, 1171).

Moreover, a close reading of Kumho makes it clear (contrary to Grove et al.’s assertions) that the relevant community is the community to which the expert belongs, namely the psychology community both academic and clinical. Referring to the objective of the gate-keeping requirement, the Kumho court opines: “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” [italics added] (Kumho Tire Co. v. Carmichael, 1999, p. 1176). Clearly, when applied to
clinical psychology, “the practice of an expert in the relevant field” means just that—what working clinical psychologists, applying both relevant research and their own professional experience, do in their everyday practice.

The admissibility standards of Daubert/Kumho were never intended to be used to exclude clinical or experience-based testimony as Grove et al. suppose. While in their other publications, some of the authors have taken a dim view of experience-based opinions (e.g., Garb, 1989), Daubert originated as an attempt to apply FRE 702, which is by no means purely about scientific knowledge. According to FRE 702, an expert “may be qualified by scientific, technical, or other specialized knowledge.” The Kumho court plainly stated, “No one denies that an expert might draw a conclusion from a set of observations based on extensive and specialized experience,” (Kumho Tire Co. v. Carmichael, 1999, p. 1178). Nor does the methodology used by a clinician need to be universally accepted. In further interpreting Daubert, Chief Judge Posner held “...that the opinion evidence of reputable scientists is admissible in evidence in a federal trial even if the particular methods they used in arriving at their opinion are not yet accepted as canonical in their branch of the scientific community” (p. 234).

Moreover, it is not clear that testimony exclusively from academic ivory towers has any real advantage in meeting the Daubert/Kumho tests of reliability and relevance. As Krauss and Sales (1999) observe: “The present standard encourages lawyers and experts to present science that may be relevant to legal issues at a broad level, but does not meet the fit requirement of Daubert when considered in light of the specific issue that the expert is supposed to address ... An irony in this result is that scientists, touting Daubert testimony over clinical opinion testimony, may be offering testimony that is no more relevant, reliable, and helpful than that which a clinician could have proffered” (p. 95).

Daubert/Kumho requires a balance between the poles of reliability and relevance, such that neither is sufficient for purposes of admissibility. Thus, clinical personality diagnoses are rarely challenged, whereas research-based testimony about the cognitive limitations applying to eyewitness testimony has often been excluded. This is not because the former are more reliable than the latter, but because nomothetic findings obtained from general research do not necessarily apply to the particular eyewitness in question and because it is up to the jury to decide matters of credibility (Slobogin, 1999). Indeed, academic experts on psychological test construction and research who have never actually been trained in the use of the Rorschach and have never applied it in an actual case might arguably face exclusion even as rebuttal experts under Daubert/Kumho.

Grove et al. (citing Wood, Nezworski, & Stejskal, 1996) contend that a psychological test should never be accepted for forensic use until every component variable and score and every associated interpretive principle has been independently validated by appropriately skeptical researchers. The upshot of this absurdly strict methodological principle would be that no psychological test more sophisticated than a checklist used to predict symptoms a patient would describe if asked would be likely to be admitted in court. Paradoxically, psychological experts would be restricted, in effect, to relying entirely on the hearsay (i.e., the ipse dixit) of the patients they were examining, certainly a dubious procedure when the examination is occurring (as it often is, in legal matters) in an adversarial context. Moreover, such an approach bears no reasonable relation to how most
tests are actually used and applied in everyday clinical practice. One might reasonably argue that research based on such principles should be ruled irrelevant by a court in a Daubert hearing. The kind of research that would meet FRE 702’s “helpfulness” criterion is clinically relevant research (see Bornstein, 2001; McGrath, 2001).

If Grove et al.’s approach to evaluating test validity were taken seriously by the courts, the ordinary diagnostic and descriptive procedures of psychologists and virtually all medical experts, would be subject to probable exclusion. The ordinary process of medical assessment involves an experience-based combination of signs, symptoms, and test results, leading to probabilistic conclusions. Both psychological assessment and medical diagnosis have been described as a process of “iterative hypothesis testing.” That is, initial clinical impressions are modified and refined as new information is considered. According to the Federal Judicial Center Reference Guide on Medical Testimony (Henefin, Kipen, & Poulter, 2000):

Medical diagnosis is not an exact science . . . Physicians make probabilistic judgments on a day-to-day basis, even when they can supplement a patient’s history and physical with the results of extensive laboratory tests (p. 465) . . . Diagnostic reasoning is usually more complex than the examples given because it is simultaneously based on multiple signs, symptoms, and test results . . . These findings are not all truly independent of one another, thus preventing straightforward addition of the probabilities as in a Bayesian model . . . Doctors combine probabilities of disease (prevalence) with their knowledge of the frequency of signs and symptoms in a given disease and competing diseases to progressively modify and ultimately arrive at their view of the likelihood of the disease under consideration.” (p. 467).

The process of arriving at a medical diagnosis as described above is structurally quite similar to that of psychological assessment. It differs dramatically from the static zero-order correlation sign-criterion relationships which seem to preoccupy Grove et al. To take the Rorschach specifically, the Comprehensive System provides guidelines (e.g., systematically reviewed interpretive hypotheses) to make personality assessment as systematic and comprehensive as possible (Exner, 2001b; Weiner, 1998).

We turn briefly to consider some of the specific criticisms Grove et al. offer concerning the Comprehensive System. The authors claim that there are serious flaws in the nonpatient norms of the Comprehensive System. Their conclusion is supported by data collected on U.S. nonpatient adults (Shaffer, Erdberg, & Haroian, 1999) and U.S. nonpatient children (Hamel, Shaffer, & Erdberg, 2000) showing that many individuals in these samples fall outside the normative range of the Comprehensive System norms on numerous variables. Grove et al. concluded that the Comprehensive System incorrectly brands nonpatients as “pathological.” However, it should be noted that, compared with the studies on which Rorschach norms are based (Exner, 1993), these studies: (a) used much smaller samples (123 in Shaffer et al. and 100 in Hamel et al.), (b) collected data in a less random fashion, (c) recruited participants from a much more limited geographic base, and (d) used less experienced examiners to administer the tests. Many methodologists might point out that Exner’s much larger sample (n = 600),
selected from a broad range of sources with much wider geographic distribution, and more experienced examiners (no students) is a sample more likely to be representative of adequate psychological functioning. If Grove et al. observed that the Shaffer et al. (1999) and Hamel et al. (2000) results gave reason to question what the Comprehensive System norms represent, we would be closer to agreement with them. However, the inference that the Shaffer et al. and Hamel et al. results invalidate the Exner norms is not methodologically sound.

Another reason given by Grove et al. for doubting the Comprehensive System norms is the recent discovery that 221 cases were accidentally entered twice in the calculation of the nonpatient norms last published in the fourth edition of the Comprehensive System workbook. This discovery had not been made when we submitted our original reply to Grove and Barden. We, of course, share Grove et al.’s consternation that such an error occurred and went undiscovered as long as it did. However, we do not share their conclusion that the occurrence of this error invalidates the current Comprehensive System norms. In correcting the error, Exner eliminated the duplications and added 121 nonpatient protocols with the same sampling and administration procedures to comprise the current sample of 600 nonduplicated cases. When the means and standard deviations of the previous norms are compared with the current statistics, the differences are negligible. Assessment psychologists using the Comprehensive System do not have to adjust their interpretive guidelines. Grove et al. actually acknowledge that “...the normative samples are highly similar” (p. 219).

Grove et al. conclude their criticisms of the Comprehensive System norms on the grounds that they are outdated and were collected when the System used somewhat different scoring criteria. We, and Exner, share this concern. Fortunately, a contemporary nonpatient sample is being collected. As of September 2001, 175 cases had been recorded. In a preliminary summary, Exner (2002) reported results strikingly similar to the original norms. Preliminary results from a contemporary multinational collaborative study in 12 countries with 2,250 respondents also suggested that many of the original Comprehensive System norms appear to be valid (Erdberg & Shaffer, 1999).

Even if the Rorschach Comprehensive System might benefit from updated norms, this would hardly set it apart from other standard, widely used and accepted psychological instruments. The MMPI struggled for nearly 50 years with a normative sample drawn largely from White, rural Minnesotans. Experienced clinicians were well aware, long before the test was revised, of the slight (about .5 standard deviations) tendency of the original MMPI to pathologize test-takers from a broader population sample and adjusted their interpretations accordingly. The Wechsler–Bellevue went through four standardizations on its way to become the Wechsler Adult Intelligence Scale—III. The Halstead–Reitan (one of the most venerable and best validated instruments in neuropsychology) has yet to establish even a first adequate normative sample. None of these tests has been abandoned by psychology (or by the courts) because its norms needed to change with the times.

Grove et al. next take on Comprehensive System reliability and validity. They claim that we failed to cite studies concerning the Rorschach’s inter-rater reliability. This was not our focus since it was not an issue emphasized by Grove and Barden in their original article. However, a quick perusal of our references shows
that we referred to several studies that directly addressed this topic (Hilsenroth, Fowler, & Padawer, 1998; McCann & Dyer, 1996; Meyer, 1997a; Parker, Hanson, & Hunsley, 1998; Stricker & Gold, 1998). To this list one can add several other studies including Hunsley and Bailey (1999); McDowell and Acklin (1996); Meyer (1997b); and Meyer et al. (2001). Grove et al.’s critique of inter-rater reliability on the Rorschach focuses on one study (Acklin, McDowell, Verschell & Chan, 2000) involving a comparison of the scores of two graduate students, whereas Meyer’s (1997a) meta-analysis of 16 published studies showing excellent inter-rater reliability is given short shrift.

Grove et al. grant that at least three Rorschach variables are valid—the Rorschach Prognostic Rating Scale (RPRS), the Rorschach Oral Dependency Scale (ROD), and the Schizophrenia Index (SCZI). They then argue that “. . . these variables are plainly irrelevant to the current debate, insofar as none are now part of the RCS” (p. 223). This statement is technically but not practically true. The RPRS, a variable developed by Klopfer (Klopfer, Kirkner, Wisham, & Baker, 1951) and recently reviewed via meta-analysis by Meyer and Handler (1997, 2000), consists of the following variables: human movement, animal movement, inanimate movement, texture, vista, shading use problems, color, color use problems, and form quality. All these variables are included in the Comprehensive System except shading use problems and color use problems. The ROD, although not currently in the Comprehensive System, will be added in the near future because the Rorschach Research Council, commissioned by Exner, has nearly completed the kind of careful evaluation of the ROD applied to all current variables, and determined that the ROD meets the standards for reliability and validity necessary for inclusion in the Comprehensive System. Finally, while the SCZI no longer exists in the Comprehensive System, a modified version with many of the same component variables, now referred to as the Perceptual Thinking Index (PTI), has been in the Comprehensive System for almost two years.

Grove et al. correctly state that the SCZI was one of the few Comprehensive System variables that corresponded to DSM–IV diagnosis and they deplore its “loss” even though they observe that it too often was found to be significant in individuals who were not schizophrenic. However, recently presented data on the validity of the PTI for predicting schizophrenia yields fewer false positives and false negatives than the old SCZI (Exner, 2001a). Grove et al.’s unfamiliarity with the current status of the RPRS, ROD, and SCZI illustrates their lack of familiarity with the Comprehensive System. Obviously, they do not have the basic, up-to-date knowledge of the Comprehensive System ordinarily expected of an assessment psychologist who is adequately prepared to use and evaluate the method in its current form.

In their discussion of the DEPI, HEV, and S-CON, Grove et al. have accused us of failing to review the entire literature. Because we only referred to these variables as examples of the few criterion-related Rorschach predictors for which known error rates have been established, we did not consider it necessary to review the entire body of literature. Now asked for our opinion, we reiterate that the Rorschach should not be used as a simple predictor of DSM–IV diagnoses. When used as we have advocated, for purposes of personality description, the DEPI does indeed appear to predict certain key elements of vulnerability to
depression, HEV (a component of the well-validated HEI) issues of personal identity and object relations, and S-CON the likelihood of suicide attempts in samples of people otherwise at risk (e.g., Fowler, Piers, Hilsenroth, Holdwick, & Padawer, 2001; Ganellen, 2001; Jansak, 1997; Perry & Braff, 1994; Renteria & Meyer, 2001; Exner & Wylie, 1977).

While Grove et al. imply that the meta-analyses supporting the Rorschach are based on a handful of studies, in fact, Parker et al. (1988) analyzed a total of 411 studies and Hiller et al.’s (1999) meta-analysis used 30 MMPI studies and 30 Rorschach studies over the past 2 decades, including 2,276 Rorschach protocols. Grove et al. acknowledge that the average validity coefficients for the Rorschach have generally been approximately .30, a figure that Cohen (1988) observed may be near the maximum limit for relating personality measures to real-life criteria. The associated “error rates” are well within the same range as those found for other tests used every day in the consulting room and in the courthouse, such as the MMPI and the WAIS (Parker, Hanson, & Hunsley, 1988; Atkinson & Cyr, 1984; Hiller et al., 1999; Meyer & Archer, 2001; Rosenthal, Hiller, Bornstein, Berry, & Brunell-Neuleib, 2001). Grove et al.’s arguments regarding unpublished studies (the so-called “file drawer” problem) are disingenuous, inasmuch as these problems have already been thoroughly addressed (Hiller et al., 1999; Meyer & Archer, 2001; Rosenthal et al., 2001).

If Grove et al. wish to argue that the results of global meta-analyses do not provide compelling support for the use of the Rorschach, they must apply their skepticism equally to the field of psychological assessment as a whole. Following a detailed review of the relevant meta-analytic literature and an expanded analysis of Parker et al.’s data set, Meyer and Archer (2001) recently concluded:

"Across journals, decades, aggregation procedures, predictor scales criteria, and participants, reasonable hypotheses for the vast array of Rorschach, MMPI, and WAIS scales that have been empirically tested produce convincing evidence for their construct validity. In turn, this knowledge provides a global foundation that documents the hard science behind the Rorschach (p. 491)."

We agree with Grove et al. that there are useful and important things that meta-analyses do not tell us, and have already argued that “error rate” is not a particularly useful concept in understanding the specific strengths and limitations of the Rorschach. As we took pains to explain in our earlier article (2002), we believe that measuring global error rates is a poor way to judge a personality test for a specific purpose, since the concept of error rates makes more sense in evaluating criterion validity than construct validity. However, if a particular court wishes to consider it in response to a global attack on the test such as that by Grove and Barden’s, these meta-analytic data are relevant (see Tenopyr, 1999).

We have explained our views on incremental validity and the Rorschach at some length. The interested reader can find a lively and much more detailed discussion of this issue in the special series in Psychological Assessment and the Journal of Personality Assessment. The important point for our present focus on the admissibility of Rorschach-based clinical testimony is that incremental validity is a question that should concern judges and courts very little. Even if everything the Rorschach can do could be done as well or better by some other instrument (a point which has never been demonstrated by even the Rorschach’s
most implacable critics), this is not an argument that supports the exclusion of Rorschach testimony. For purposes of forensic testimony, it is standard practice for psychologists to use multiple methods of data collection (American Psychological Association, 1994) with the goal of using results gathered by the various methods to confirm each other. Technical arguments over whether to use three instruments when two could do just as well are of little consequence to the legal system. As for the now rather tired criticisms of the work of Gacono, Meloy (e.g., Meloy, Hansen, & Weiner, 1997), and others regarding psychopathy, we believe they have been adequately addressed (Gacono, Loving, & Bodholt, 2001; but see Wood, Lilienfeld, Nezworski, & Garb, 2001).

In one of the few arguments directly addressing Daubert/Joiner/Kumho standards, Grove et al. reiterate that because many studies attesting to the reliability and validity of the Comprehensive System appear in the Journal of Personality Assessment, a “speciality journal,” the Comprehensive System has not been subjected to appropriate peer review. They contend that because the Journal of Personality Assessment was previously entitled the Rorschach Research Exchange, the Journal of Projective Techniques, and the Journal of Projective Techniques and Personality Assessment, it has a “lengthy history of close association with this technique” (p. 224). This argument is curious, because many of the studies cited by Grove et al. and their colleagues as having detected flaws in the Comprehensive System were published in the Journal of Personality Assessment—evidence, one would think, of the objectivity of the journal and its respectful recognition of divergent perspectives within the field of assessment.

Further disparaging the Journal of Personality Assessment, Grove et al. (correctly) observe that every editor (Bruno Klopfer, Walter Klopfer, Irving Weiner, Bill Kinder, and Greg Meyer) was an assessment psychologist who studied and evaluated the Rorschach. Can we assume from this argument that since James Butcher, the chief architect of the revised MMPI-2, is editor of Psychological Assessment, all studies of the MMPI-2 in that journal are suspect?

Grove et al. end their rejoinder by deploring untrained and incompetent use of the Rorschach Comprehensive System in forensic practice and correctly presume that we do as well. But this issue is irrelevant to the general admissibility of the Comprehensive System for expert testimony. Should brain surgery never be performed because somebody might make a mess of it? Condemnation of the misuse of the Comprehensive System does not meaningfully address whether the Comprehensive System, used properly by an expert, meets the standards developed in Daubert and its progeny.

Grove et al. seem to be trying to gain in the legal system what they cannot accomplish in the scientific and professional journals. The Rorschach is a test that has consistently proved, when properly used as part of a broader psychological assessment, to have accuracy comparable to the two other most widely used tests in the field of psychology, the MMPI and the WAIS (Atkinson & Cyr, 1984; Hiller et al., 1999; Meyer, 2000; Parker et al., 1988; Rosenthal et al., 2001), to compare favorably in predictive validity with many standard medical laboratory tests (Meyer, 2000), and to have been accepted without question in clinical training institutions (Clemence & Handler, 2001) and courtrooms (Meloy, Hansen, & Weiner, 1997; Weiner, Exner, & Sciara, 1996). However, Grove et al. argue that this test must meet the special requirement that all of its current critics
be silenced before it sees the light of day in an American courtroom. Far from the liberalization of the “general acceptance” standard intended by Frye v. United States (de-emphasized but still considered in Daubert), Grove et al. urge a radical retrenchment that amounts to a “universal acceptance” standard (or worse, a demand that all clinical testimony be reviewed prior to admission into evidence academics who may have seldom, if ever, practiced personality assessment).

Grove et al. confuse the current limited debates over various technical issues with the Rorschach (in which they have served among the primary adversaries of the status quo) with evidence that the technique is still not yet ready (if it will ever be) for the courtroom. Ostensibly to protect scientifically naïve judges from being snookered by fast-talking clinicians, they show every indication of wishing to set themselves up as the arbiters of standards for psychological testing. We disagree.

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


Meyer, G. J., & Handler, L. (1997). The ability of the Rorschach to predict subsequent


