Psychology, Public Policy, and Law
1999, Vol. 5, No. 1, 224–242

1076-8971/99/$5.00 DOI: 10.1037//1076-8971.5.1.224

PROTECTING THE INTEGRITY-OF THE LEGAL SYSTEM
The Admissibility of Testimony From Mental Health Experts Under *Daubert/Kumho* Analyses

William M. Grove
University of Minnesota at Minneapolis
St. Paul

R. Christopher Barden
National Association for Consumer
Protection in Mental Health Practices,
North Salt Lake, Utah

The authors discussed to what degree testimony from social science and mental health experts (psychologists, psychiatrists, social workers, therapists, others) meets admissibility requirements expressed by the U.S. Supreme Court in *Daubert* (1993), *Joiner* (*General Electric Co. v. Joiner,* 1997) and the recent *Kumho* (1999) decision. They reviewed data on *Daubert/Kumho* indicia of reliability using 2 exemplar areas of mental health testimony: psychodiagnostic assessment by means of the Rorschach and other "projective" assessment techniques and the diagnoses of posttraumatic stress disorder and multiple personality disorder (dissociative identity disorder). They concluded that some testimony offered by mental health professionals relating to these concepts should not survive scrutiny under the framework of *Daubert, Joiner,* and *Kumho.*

Prior to the ruling of the U.S. Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993), testimony from mental health and social science experts was largely unregulated by the legal system. The *Frye* (*Frye v. United States,* 1923) standard had been in place for decades (Gianelli, 1980), requiring that to be admissible, the scientific bases of testimony must be "generally accepted" in the "field" to which they belong. This is a very lenient standard; experts can always be found who will swear that a theory is "generally accepted." Under *Frye,* the expert is not required to substantiate the scientific soundness of the theory by reference to proper research documenting other hallmarks of a reliable theory, such as the theory's survival of Popperian risky tests, survival of peer review, or calculable error rates. Moreover, "general acceptance" itself is usually established by the expert's say-so (subject to the finder of fact's judgment about the expert's credibility); citation of survey studies that document such acceptance are usually not required. Hence, testimony by mental health professionals regarding all sorts of controversial theories and methods has very often been admitted under *Frye.*

The 1993 *Daubert* ruling of the U.S. Supreme Court changed this unfortunate situation and heightened interest in, and concern about, expert testimony based on "junk science." In *Daubert,* the U.S. Supreme Court ruled that scientific expert

William M. Grove, Department of Psychology, University of Minnesota; R. Christopher Barden, National Association for Consumer Protection in Mental Health Practices, North Salt Lake, Utah.
We thank many colleagues for reviewing a draft of this article.
Correspondence concerning this article should be addressed to William M. Grove, Department of Psychology, University of Minnesota, N218 Elliott Hall, 75 East River Road, Minneapolis, Minnesota 55455-0344. Electronic mail may be sent to william.m.grove-1@tc.umn.edu.

testimony is admissible only if it is both relevant and reliable. The Court further held that the Federal Rules of Evidence "assign to the trial judge the task of ensuring that an expert's testimony both rests on a reliable foundation and is relevant to the task at hand" (*Daubert,* 1993, p. 597). In addition, the Court discussed specific factors useful in determining the reliability of a scientific "theory or technique" (*Daubert,* 1993, pp. 593–594), generally following the philosophy of science of Sir Karl Popper (1959). We discuss these factors in the sections that follow.

Clarifying the review process enunciated in *Daubert,* the U.S. Supreme Court later ruled that the court of appeals must apply an abuse-of-discretion standard when it reviews the trial court's decision to admit or exclude expert testimony (*General Electric Co. v. Joiner,* 1997). The *Joiner* ruling further emphasized the responsibility of the trial court to fulfill a mandatory gatekeeper role—the duty to exclude unreliable expert testimony—assigned by the *Daubert* ruling.

In a March 1999 decision with enormous importance for the regulation of the testimony of mental health professionals, the U.S. Supreme Court expanded the *Daubert* analysis to the testimony of essentially all expert witnesses (*Kumho Tire Co., Ltd., et al. v. Carmichael et al.,* 1999). In *Kumho,* the defendant tire company moved to exclude the plaintiff's expert testimony on the ground that his methodology failed to satisfy Federal Rule of Evidence 702, which states: "If scientific, technical, or other specialized knowledge will assist the trier of fact, a witness qualified as an expert may testify thereto in the form of an opinion." Granting the motion to exclude the expert testimony in question and entering summary judgment for the defendants, the District Court acknowledged that it was acting as a reliability gatekeeper as required by *Daubert.*

Opining that *Daubert* was limited to "scientific" testimony, the U.S. 11th Circuit Court of Appeals held that the *Daubert* factors did not apply to the expert's testimony, which it attempted to distinguish as "skill- or experience-based." The U.S. Supreme Court overturned the 11th Circuit's interpretation and reinstated the District Court opinion by holding that the *Daubert* factors may apply to the testimony of engineers and other experts who are not claiming a basis for their testimony in rigorous scientific research and peer reviewed publications (*Kumho,* 1999, pp. 7–13). In an opinion written by Justice Breyer—a justice highly trained in methodology and philosophy of science issues—the Court ruled that the *Daubert* "gate keeping" obligation applies to all expert testimony. The Court held that judges would find it hard, if not impossible, to perform this function while distinguishing between "scientific" knowledge and "technical" or "other specialized" knowledge and that there is no clear line dividing the one from the others and no convincing need to make such distinctions (*Kumho,* 1999, pp. 7–9). Prior to Kumho, unscientific experts could testify if they showed the "same level of intellectual rigor that characterizes the practice of an expert in the relevant field." This loophole was noted by Justice Breyer, writing for the Court: "As the U.S. Supreme Court ruled in *Joiner,* 'nothing in either *Daubert* or the Federal Rules of Evidence requires a district court to admit opinion evidence that is connected to existing data only by the *ipse dixit* of the expert' " (*Kumho,* 1999, p. 146).

In the *Daubert* (1993), *Joiner* (*General Electric Co. v. Joiner,* 1997), and *Kumho* (1999) cases, the U.S. Supreme Court has begun the long-overdue process of educating legal professionals in the essential, minimal characteristics of science.

Six factors of scientific analysis, or indicia of testimonial reliability, can be distinguished in *Daubert:*

1. Is the proposed theory, on which the testimony is to be based, testable (falsified)?
2. Has the proposed theory been tested using valid and reliable procedures and with positive results?
3. Has the theory been subjected to peer review?
4. What is the known or potential error rate of the scientific theory or technique?
5. What standards, controlling the technique's operation, maximize its validity?
6. Has the theory been generally accepted as valid in the relevant scientific community?

As other contributors to this issue have pointed out (e.g., Krauss & Sales, 1999, this issue), these features were not enumerated as an exhaustive list. Furthermore, the *Daubert* Court did not require trial judges to combine these factors algorithmically in deciding on admissibility, nor did they assign weights to the factors. Hence, it has been left to case law to clarify the proper application of *Daubert.* The reader is referred to other articles where these features are discussed from a legal point of view (e.g., Lipton, 1999, this issue; Schopp, Scalora, & Pearce, 1999, this issue) and pertinent cases are analyzed. Bersoff, Glass, Dodds, Eckl and Peters (1999) have compiled a list of federal appellate *Daubert* cases.

Daubert admissibility hinged on acceptability of a theory or methodology, and not on conclusions drawn from it. However, the Joiner Court held that "[c]onclusions and methodology are not entirely distinct from one another; when the analytical gap between the data and the opinion proffered is simply too great [the evidence may be excluded]" (*General Electric v. Joiner,* 1997). Therefore, a seventh factor can be added to those above: Do the expert's conclusions reasonably follow from applying the theory to this case?

We consider two exemplar areas of expert mental health testimony that show how we go about analyzing *Daubert–Joiner–Kumho* factors: the Rorschach test and controversial diagnoses (e.g., posttraumatic stress disorder [PTSD], and multiple personality disorder, now known as *dissociative identity disorder* [MPD/DID]). We argue that the use of the Rorschach test for psychodiagnosis and personality description, and common courtroom testimony about PTSD and MPD diagnoses, should be found inadmissible. We end the article by considering how generalizable our examples are to forensic mental health testimony in general.

Rorschach Testing for Diagnosis and Personality Description

Of the many extant projective tests, none has as much supporting research as the Rorschach. If it can be shown that testimony based on the Rorschach, when used for diagnosis of mental disorders and personality description, is inadmissible under *Daubert,* it likely follows that no other projective technique is admissible, either.

We prefer to deal solely with *The Rorschach: A Comprehensive System* (TRACS), Exner's (1993) method of administration, scoring, and interpretation. This method is the most researched of several Rorschach "schools," and confining attention to TRACS would show the Rorschach to best advantage. However, sole attention to TRACS is not feasible for two reasons. First, many studies do not use TRACS or do not use it carefully, and so generalization to TRACS-based interpretations may be hazardous. Second, many clinicians do not use TRACS (Exner, 1980), and generalizing from TRACS-based studies may not accurately estimate the accuracy of non-TRACS interpretations.

Daubert concerns scientific theories, but a mental test is not a theory. Although early Rorschach workers proposed the "projective hypothesis" as a theoretical basis for the Rorschach, use of the Rorschach today seems instead to rely on the following less controversial theory: the Rorschach, when administered, scored, and interpreted in a specific manner, yields accurate diagnoses of psychological disorders or accurate personality descriptions. (We bypass the occasional objection that the Rorschach is not a test, because in the present context this is a distinction without a difference.)

1. Is this technique testable? Yes. It is relatively straightforward to determine whether an instrument yields scores that are valid, in the sense of notably correlating with external criteria including psychiatric diagnoses.

2. Has this technique been tested? Yes. This is true for two important aspects of validity. The first concerns zero-order validity correlations between Rorschach scores and various criteria, such as psychiatric diagnoses or non-Rorschach personality measures. The second is incremental validity, which is the ability of a test to add to diagnostic accuracy or personality description, when added to other commonly obtained information (e.g., chart data, interviews, Minnesota Multiphase Personality Inventory [MMPI–2]; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989).

The interpretation of zero-order validity studies is currently controversial; Rorschach proponents see the validity more optimistically than do skeptics. We read the literature as suggesting that typical Rorschach scores have low validity, in the .2–.3 range, somewhat lower than the MMPI (Garb, Florio, & Grove, 1998; Garb, Wood, Nezworski, Grove, & Stejskal, in press). Incremental validity studies, on the other hand, quite consistently show zero to negative validity for the Rorschach (Garb, 1985).

One could argue that if the Rorschach has any zero-order validity, it passes this *Daubert* test. However, it is more important to remember that without incremental validity, the Rorschach—on average—adds nothing to diagnostic evaluations. Indeed, the Rorschach can be expected to detract from testimonial accuracy, if it has no incremental validity. This is because it enhances confidence without enhancing validity. Psychologists (and other clinicians) often show increased confidence in their judgments when they have more information rather than less; but their judgments often do not grow more accurate as more information accures (Garb, 1985; Oskamp, 1965). Unwarranted but confidently expressed opinions from a testifying expert are more likely to be prejudicial than probative, and hence should not be admitted.

3. Is the Rorschach generally accepted as valid for diagnosis and personality description? No; it is quite controversial among personality assessment researchers and always has been. Courts are encouraged to analyze years of negative Rorschach reviews in authoritative sources, including Buros Institute's *Thirteenth*

Mental Measurements Yearbook (Impara & Plake, 1988; see e.g., Cronbach, 1949; Jensen, 1965; Nezworski & Wood, 1995; Wood, Nezworski, & Stejskal, 1996).

4. What is the error rate associated with Rorschach interpretations and diagnoses? This question cannot be succinctly answered, because accuracy varies across applications. Unquestionably, however, the Rorschach sometimes yields highly inaccurate diagnoses. For example, in a classic pre-TRACS study, Little and Schneidman (1959) suggested that the error rate can be 100%; in their report, the modal diagnosis given to psychiatrically undisturbed individuals by Rorschach experts was one of psychosis. These were standard Rorschach protocols, taken verbatim and interpreted by renowned experts (e.g., Zygmunt Piotrowski), but without benefit of direct comparison to normative data and without use of statistical prediction from scores. Using a similar expert–judge design but also with no explicit normative comparison or statistical prediction based on scores, Albert, Fox, and Kahn (1980) compared malingerers and psychotic individuals. They found 52% diagnostic errors among psychotic individuals, 46–72% errors with malingerers, and 24% with controls. In criminal forensic contexts such error rates are clearly unacceptable. For nonpsychotic diagnoses (e.g., depression), the Rorschach's validity is also quite weak (Wood, Lilienfeld, Garb & Nezworski, 1999; Wood et al., 1996).

Meta-analyses of the Rorschach (e.g., Parker, Hanson, & Hunsley, 1988) suggest typical Rorschach score validities of about .3. It is easily calculated that if a Gaussian measure has .3 validity, then individuals classified into the more abnormal group represent classification errors over 38% of the time (when the abnormal group has a relative frequency of 50%). This error rate climbs as the abnormal group's frequency goes down, being over half when the frequency is 10%.

The error rate for personality descriptions based on the Rorschach is likewise very substantial. If the validity of a Rorschach score is .3, then the error rate of a typical personality trait inference is 91% as great as that for completely random Rorschach interpretations. (In real life, the typical error rate almost surely exceeds this figure, because the typical validity figure of .3 stems from studies that statistically compare groups. Clinicians generally do not predict nearly as well as do statistical formulae; Grove et al., in press.) In our opinion, the adjudicated rights of citizens should not turn on such an error-prone way of obtaining diagnoses and personality descriptions.

5. Has the validity of the Rorschach been subjected to peer review? Yes, no, and maybe. The assessment community has been surprised to learn in recent years that many of Exner's studies cited in his 1993 volume as supporting TRACS, were apparently never peer reviewed; in fact, many do not even exist as actual written reports, and their data are not readily available to other investigators (Wood et al., 1996). Furthermore, pro-Rorschach studies often appear in a specialty journal, the *Journal of Personality Assessment* (originally called the *Rorschach Research Exchange*), which may potentially relate to quality of publication. Finally, the effectiveness of peer review in this area has been questioned, given notable problems with clarity and accuracy of some published Rorschach studies (Wood, Nezworski, Stejskal, Garven, & West, 1999).

6. From the increasing popularity of Exner's scoring and interpretation system, it might seem that using TRACS is a well-accepted way of enhancing the Rorschach's accuracy. However, this is not so. Hiller, Rosenthal, Bornstein, Berry,

and Brunell-Neuleib (1999) claimed to find evidence for superior validity of TRACS when compared to other systems. However, they reach this sanguine conclusion by misinterpreting a nonsignificant difference ($p \lesssim .175$, two-tailed) as favoring TRACS (Garb et al., in press).

Based on information from Rorschach studies, it would seem that there is a way to maximize accuracy: confine Rorschach use to one or a few well-validated scores (e.g., F + % or X + %) that have been shown to reliably predict a diagnosis of psychosis. One might thereby do somewhat better than chance in making such diagnoses; but the error rate is still substantial. However, three problems prevent directly applying research error rates for this kind of problem to forensic clinicians. First, forensic experts apparently do not rely on statistical predictions from normed Rorschach scores; instead, they clinically combine information and do not explicitly reference good normative data. In addition, experts seldom testify based on a single score; mixing valid with invalid (or less valid) scores can easily "wash out" all the valid diagnostic information in a test protocol, owing to the deficiencies of clinical as opposed to statistical prediction (Grove et al., in press). Finally, in the typical clinical forensic situation, the Rorschach data come into a picture with a lot of other data on hand (life history data, interview information, other psychological test scores); hence, the error rates from Rorschach studies, which are essentially always based on zero-order validity, cannot be used to infer error rates for this kind of incremental validity situation.

In contrast to our analysis of Rorschach admissibility, McCann (1998) published a more sanguine review, but it antedates *Kumho.* Guidelines for ostensibly appropriate use of the Rorschach in court were offered by Meloy (1991). Weiner, Exner, and Sciara (1996) noted that the Rorschach has seldom been excluded in court (they noted only six challenges, one of them successful, out of 7,934 cases using Rorschach testimony).

The above is consistent with our experience that very few attorneys conduct rigorous *Daubert* (or even *Frye*) hearings to exclude unreliable, junk science testimony. We find the use of the Rorschach to be quite problematic under *Daubert/Kumho.* We also believe that the analysis in McCann (1998, pp. 133–140) is seriously flawed. McCann did not explain how Rorschach-generated impressions of a client are falsifiable, a fundamental *Daubert* criterion. He also misinterpreted the continuing controversy over this test as somehow constituting evidence for its general acceptance. (He does not note that a controversy can continue long after most disputants have reached a negative opinion.) McCann also erred in relying on Exner's (1993) self-cited, unpublished reports, which as discussed above are often not peer-reviewed or even physically extant manuscripts. McCann likewise relied on unreplicated and unrepresentative accuracy figures for selected Rorschach variables in concluding that the Rorschach has acceptable error rates under *Daubert.* He curiously did not rest his error rates on the then-best available review of Rorschach accuracy: Parker et al.'s (1988) meta-analysis. This review yielded two average Rorschach validities: .42 for "convergent" validity studies and .07 for other studies. Skeptical of this report, Garb et al. (1998) reanalyzed Parker et al.'s data and found an average validity of .3, which we used here for error rate calculations (above). We suggest that experts and courts not rely on McCann's analysis as definitive.

Controversial Diagnoses (PTSD and MPD/DID)

We consider just two diagnoses here: PTSD and MPD, now called DID in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM–IV;* American Psychiatric Association, 1994). First, we note that PTSD and MPD are *labels,* not theories. This is a critical distinction—some experts have misled courts to believe falsely that the existence of a diagnostic label in *DSM–IV* somehow proves general acceptance of the existence of the described disorder as well as acceptance of proposed causal mechanisms for the etiology of such a disorder. This is a very serious error of logic and method. The *DSM–IV* is simply an agreed upon set of terms and descriptions—a catalog. It does not provide, and was not intended to provide, documentation of the general acceptance of the existence of disorders. Furthermore, the *DSM–IV* is not in any way documentation of general acceptance of the etiology (cause) of a disorder. To clarify with an example: The word *unicorn* is in the dictionary and we all agree on the concept and description of a unicorn, but this surely does not document the existence of unicorns.

DSM labels, considered only as labels, may not be generally accepted, even though they make it into the manual (just as the Volstead Act [National Prohibition Act, 1919] stayed on the books years after public support for it had eroded). With regard to the *DSM,* the manual's cautionary statement describes the categories and criteria as a "consensus" position. However, this must not be understood to imply any general polling to establish consensus. Indeed, the manual's introductory explanation of the *DSM* revision process makes clear that no polling was involved. Instead, specialty subcommittee (e.g., there was a subcommittee on dissociative disorders) members were assigned to review aspects of the literature relating to certain categories, using subjective methods (i.e., no requirement for meta-analysis) and following a common format. Analyses of existing data sets were sometimes undertaken. These reviews were then critiqued by others, chosen by Task Force members. Subcommittees then voted for revised labels (or criteria), and the Task Force voted whether to accept a subcommittee's recommendation. This procedure is defensible, but it is not completely explicit, repeatable, or tied to polling representative samples of scientists.

McHugh (1998) explained the distinction between *DSM* category inclusion and category validation in a *Daubert* hearing:

> Q: . . . is the fact that the term "dissociative amnesia" is found in the *DSM* any evidence at all of the scientific validity of the condition?
> A: No, it's not. *DSM–IV* is an attempt to, as *DSM–III* was, is an attempt to develop reliability among psychiatrists about what they are observing. It is not intended to say that they are claims for the existence of a particular condition as confirmed by its enclosure within *DSM–IV.* It's a question of reliability versus validity. (McHugh, 1998, pp. 530–531)
>
> [P]sychiatry is the only discipline that runs itself on a catalog, and the reason it did that was that back in the 19—late 1960s and early 1970s, it was discovered that they couldn't do much research in psychiatry because they couldn't get agreement on what they were observing, let alone what they were calling things. So a patient with a set of symptoms called schizophrenia in Baltimore was called manic depressive in London and demoralized in San Diego. It was the decision of the American Psychiatric Association and of the American psychiatric community that we should build a catalog of reliability, not of validity, so that we could do research.

> So that when we said these patients are showing these symptoms that we are going to call Dissociative Disorder or schizophrenia, we could have a common language. After that, we could discover whether they had, in fact, anything like dementia, or whether they were behaving in a particular way for other reasons. So when I'm saying this (DSM) is a book of reliability, I'm saying this gives you a code that whereby you can describe people whom you say will be given—will satisfy the criteria in this book [DSM] for Dissociative Disorder, but whether Dissociative Disorder exists in itself is not proven or even claimed in this book [DSM]. (McHugh, 1998, p. 637)

In use, *DSM* diagnoses of MPD (or PTSD) generally are linked to accompanying theories of etiology: namely, that these psychological disorders are caused, in whole or in large part, by psychological trauma. In the case of PTSD, this includes a wide range of upsetting events; for MPD, the cause is purported to be early, severe (and almost always repressed) abuse, chiefly sexual abuse (Putnam, 1989). The following *Daubert* features are related to admissibility for PTSD.

1. The concept of PTSD, simply as a diagnostic label, is testable in principle. However, to obtain a test placing such a diagnosis at risk of falsification, one would need a benchmark for accuracy. Because many, even most, forensic PTSD evaluations now use structured interviews and claim to adhere to *DSM* criteria, validity studies seldom have a benchmark better than the forensic diagnosis itself. Spitzer (1983) proposed a method for validating diagnoses that could be useful for PTSD: the "LEAD standard," denoting longitudinal, expert, and all-data based diagnoses. Unfortunately, this useful strategy has apparently never been used to assess the concept of PTSD.

Testing the theory that PTSD is caused by exposure to trauma is also possible: Compare people who have been exposed to an objectively verified, naturally occurring traumatic event with people who have not been exposed, and find out whether the traumatized group gets PTSD more often then the unexposed group does. (Ethical principles preclude a direct experimental test.)

2. Has the causal theory of PTSD been tested? Yes, but the tests have been weak because of ethical and practical problems. Tests to date have yielded variable, mostly weak confirmations demonstrating the unreliable nature of the concept (Bowman, 1997). For example, McFarlane (1987, 1988a, 1988b, 1988c, 1988d) found that just 9% of individuals' symptoms could be accounted for by exposure to a disastrous fire, whereas Galante and Foa (1987) found no relationship between proximity to an earthquake and children's behavioral disturbances. By contrast, preexisting symptoms may predict as much as half of the variance in posttrauma symptoms (Nolen-Hoeksema & Morrow, 1991). Only for those exposed to the severest, most prolonged traumas (e.g., prisoners of war) does the rate of PTSD rise to one-half or more (Engdahl, Dikel, Eberly, & Blank, 1997). In a review of 45 studies Kendall-Tackett, Williams, and Finkelhor (1993) reported that abused children displayed more symptoms than nonabused children, with abuse accounting for 15% to 45% of the variance, but they did not analyze rates of PTSD diagnoses.

One might argue that if any PTSD-type symptoms are significantly associated with trauma, this makes PTSD testimony probative and hence admissible. However, this ignores the difference between scientific and legal concepts of causation. An expert usually has to testify as to whether a specific trauma was a "substantial factor" in creating the disorder, because causation in the "but for" sense is typically out of the question in trauma cases. As research shows that traumas usually account for only a minority of symptoms, an expert may have difficulty opining that a specific trauma is a substantial factor in causing symptoms. If "substantial factor" means "more likely results in the effect than not," then for a trauma to cause PTSD, one would need research showing that over half of those exposed to a particular type of trauma go on to develop PTSD.

Psychological causation is seldom so straightforward. Correlational studies relating traumas to symptoms often show that third variables (such as other adversity, preexisting symptoms, or coping styles) can account for much of the apparent trauma–symptom connection (Beitchman et al., 1992; Kendall-Tackett et al., 1993). If this is so, then theories like "PTSD is caused by trauma" cannot be fairly characterized as having survived strong risks of falsification. In sum, the causal connection between trauma and PTSD appears too unreliable to survive a thorough *Daubert/Kumho* analysis.

3. Is the theory of PTSD generally accepted? The label is more controversial than some (e.g., mania) but less so than others (e.g., MPD). We believe that the syndrome is generally, but by no means universally, accepted.

However, as noted above, acceptance of the label need not imply acceptance of the causal theory. PTSD positively invites misunderstanding on this score; its diagnostic requirement of a trauma event invites experts to commit the *post hoc ergo proper hoc* fallacy, assuming that the trauma caused the PTSD. The very name *posttraumatic stress disorder* seems to imply what the criteria do not state, namely causality.

If we consider an analogy from the depressive disorders, the fallacy is clarified. If *DSM–IV* erected a "reactive depression" category, with diagnostic criteria requiring a negative life event plus certain depressive symptoms, would this justify the conclusion that the negative event caused the symptoms? Of course not. In fact, negative life events have little proven relation to depression and "neurotic" ailments (Tennant, 1983; Tennant, Bebbington, & Hurry, 1981).

4. What is the known or potential error rate of a PTSD diagnosis? A relevant error rate would seem to be the frequency with which a court would err, if it equated a PTSD diagnosis with proof of traumatic causation. If symptoms are only 9% predictable from trauma exposure, as McFarlane (1988a) found, this would correspond to an error rate of over 35% in causal inferences, assuming a 50% base rate of actual trauma exposure; the error rate grows if trauma exposure is rarer than this. Because we do not know the true rate of trauma exposure suffered by plaintiffs, it is fair to say that the error rate is unknown. Such uncertain and potentially very high error rates clearly, should not survive thorough *Daubert/Kumho* scrutiny.

5. A major problem with standards in diagnosing *DSM–IV* (American Psychiatric Association, 1994) PTSD is the vagueness of its "A" criterion:

> The person has been exposed to a traumatic event in which both of the following were present: (1) the person experienced, witnessed, or was confronted with an event or events that involved actual or threatened death or serious injury, or a threat to the physical integrity of self or others; and (2) the person's response involved intense fear, helplessness, or horror. (p. 427–428)

The inclusion of the phrase "threat to the physical integrity of self or others" may allow a biased or careless "expert" to overdiagnose PTSD. Moreover, criterion

A(2) is obviously quite subjective, as are most of the symptoms of PTSD. Malingering these symptoms would not be difficult. Even though a PTSD diagnosis requires actual (not just claimed) trauma, it has been the authors' observation that many forensic clinicians do not even attempt to corroborate the trauma claim or make the PTSD diagnosis conditional on a court's determination of the accuracy of the trauma claim. This standards problem is especially striking given that, in many PTSD claims, it seems that the parties dispute all or part of the trauma claim.

For the sake of completeness, we add here the obvious fact that one cannot infer "backwards" from the existence of PTSD-type symptoms (or MPD) to the occurrence of a historical trauma. This is true for logical reasons: The cause is not deducible from its effects, because the effects have more than one possible cause. It is also true for two statistical reasons. First, the conditional probability of trauma, given the existence of symptoms, is not the same as the conditional probability of symptoms, given trauma; the latter is what PTSD research ordinarily documents. Second, the conditional probability of trauma, given symptoms, is ordinarily much lower than the conditional probability of symptoms, given trauma; this is because the frequency of specific types of trauma (e.g., sexual abuse) is much less than 50% in the relevant population.

6. With regard to peer review, many studies of the relationship between trauma and psychiatric symptoms have been published in mainstream journals (e.g., *Archives of General Psychiatry*) as well as in specialty journals (e.g., *Journal of Traumatic Stress, Child Maltreatment,* and *Child Abuse and Neglect*). As mentioned above, however, many of these studies report only weak support for the trauma–illness connection.

If PTSD diagnoses under *Daubert* are problematic, diagnoses of MPD/DID (hereinafter, MPD) are downright untrustworthy. Like PTSD, the underlying theory, not stated in the criteria, is that trauma causes the disorder. For MPD the assumed trauma is reportedly quite commonly sexual in nature, ostensibly beginning in early to middle childhood (Putnam, 1989).

In assessing the credibility of the MPD causal hypothesis, an acquaintance with the full range of theories about MPD is important. Courts must be truthfully informed that variants of the theory of MPD—crafted by central figures of MPD theory and practice—are unsupported by any credible evidence. One national leader, Corydon Hammond, has posited that MPD is often the product of "programming" by intergenerational, international Satanic cults. In a question-and-answer session with an audience of psychotherapists, he explained:

> Q: What's the difference between this kind of program and cult-type abuse and Satanic abuse in the kind of cults with the candles and the . . .
> A: This type of programming will be done in the cults with the candles and all the rest. My impression is this is simply done in people where they have great access to them or they're bloodline and their parents are in it and they can be raised in it from an early age. If they are bloodline they are the chosen generation. If not, they're expendable and they are expected to die and not get well. There will be booby traps in your way if they aren't non-bloodline people that when they get well they will kill themselves. I'll tell you just a little about that. My belief is that some people that have ritual abuse and don't have this have been ritually abused but they may be part of a non-mainstream group. The Satanism comes in the overall philosophy overriding all of this. People say, "What's the purpose of it?" My best guess is that

> the purpose of it is that they want an army of Manchurian Candidates, tens of thousands of mental robots who will do prostitution, do child pornography, smuggle drugs, engage in international arms smuggling, do snuff films, all sorts of very lucrative things and do their bidding and eventually the megalomaniacs at the top believe they'll create a Satanic order that will rule the world. (Hammond, 1992)

Similarly, Bennett Braun, former president and co-founder of the International Society for the Study of Dissociation (ISSD) and past editor of its journal *Dissociation,* has reportedly opined that MPD often results from "Satanic cult" abuse:

> [Braun] told the audience that children are often abused in day-care centers as a way of prepping them to join the cult. "You have to predispose the nervous system to this sort of behavior," he said. After the children are abused in day care, they are "then picked up in high school" and indoctrinated into the cult. The cult, he has come to understand, is networked with the Ku Klux Klan; neo-Nazi groups; the Mafia; big business; the intelligence community, including the CIA; and the military. He told the audience that he has developed twelve P's for those involved in satanic abuse: "Pimps, Pushers, Prostitutes, Physicians, Psychiatrists, Psychotherapists, Principals and teachers, Pallbearers [meaning undertakers], Public workers, Police, Politicians and judges, and Priests and clergies from all religions." (Ofshe & Watters, 1996, p. 245)

Braun was recently a co-defendant in a case alleging that he imposed his theories on a mother and her children using hypnosis and other means. Although he denied negligence, the plaintiff received a record $10.6 million settlement. Yet another former ex-president of the ISSD, Colin Ross, has written a book proposal expounding his theory that MPD cases may be caused by CIA-military brainwashing experiments (Ofshe & Watters, 1996, p. 223).

Although these theories may sound outlandish or at any rate utterly wanting in proof, professionals like Hammond, Braun, and Ross have had great influence in the MPD field. Other therapists, relying on MPD theories of these kinds (and related material published in *Dissociation* and elsewhere), have suffered multimillion-dollar jury verdicts for malpractice, and some have lost their medical licenses. These facts support the idea that such theories are not generally favorably regarded by the larger scientific or professional communities, although they have had currency in the MPD subcommunity.

Because of the sharp divergence of some of these leaders' MPD theories from more mainstream views, courts evaluating peer review for *Daubert* purposes need to understand that peer review may not always function in its normal manner (i.e., to detect and correct error). Courts may need to explore the knowledge, training, experience, and judgment of editors and peer reviewers, rather than simply assume that "peer review" is a simple, nonevaluative fact determination.

Less implausible variants of the MPD theory are also unlikely to survive *Daubert/Kumho* analysis. We now examine the six indicia, with regard to the tamer theory that MPD results from severe and prolonged child abuse, usually including sexual abuse (Putnam, 1989).

1. This hypothesis is testable in principle. A proper, prospective longitudinal study of children with known abuse would yield evidence about the rate of MPD in such children once they became adults. The use of proper controls, similar to the

traumatized children in many relevant ways other than trauma exposure, would be crucial.

2. However, the theory has not been tested in this straightforward fashion. All studies to date are retrospective and are based on self-reported abuse histories. Because MPD diagnosis and retrospective abuse histories have ill-determined error rates, existing tests of the theory are not good evidence that MPD is ever caused by severe child abuse, let alone that it is routinely so caused.

The following would be convincing evidence for the MPD-abuse theory, if shown in independent, replicated studies. Individuals with verified trauma histories are prospectively followed. They are later prospectively observed, by observers blind to the trauma histories, to show unmistakable signs of MPD (not signs shared with many other disorders, e.g., depersonalization). The only unequivocal MPD sign would seem to be obvious alter switching, because the *DSM–IV* amnesia criterion is quite vague ("amnesia for important personal information too extensive to be accounted for by ordinary forgetting"; p. 487). Concomitant variables (e.g., nonabuse adversity, family history of psychopathology, history of exposure to trauma-centric psychotherapies, or therapy with MPD advocates) would have to be convincingly ruled out as competing causes of MPD symptoms. There are no such published, peer-reviewed studies. Beitchman and colleagues (1992) reviewed the literature and concluded there was insufficient evidence of a link between child abuse and MPD.

3. Some studies of MPD/DID are peer reviewed. Much of the material is not peer-reviewed or is reviewed less stringently (books, conference presentations). Material skeptical of MPD is more likely to be published in mainstream and general journals (e.g., *British Journal of Psychiatry*), whereas much of the pro-MPD material appears in specialty journals (e.g., *Dissociation*).

4. What is required for a correct inference that abuse underlies a case of MPD? Consider whence the inference arises. From writings of MPD advocates as well as our own review of cases, it appears that abuse is inferred from "recovered" (i.e., ostensibly formerly "repressed" or "dissociated," then recalled) memories. Hence, the causal inference is usually no stronger than the evidence for "repression" or "dissociation" of trauma memories. This evidence is quite feeble (Pope & Hudson, 1995; Pope, Hudson, Bodkin, & Oliva, 1998; Pope, Oliva, & Hudson, 1999), and courts applying *Frye* or *Daubert* have very seldom held that this concept is admissible (*Barrett v. Hyldburg,* 1996; *Blackowiak v. Kemp,* 1996; *Borawick v. Shay,* 1995; *Carlson v. Humenansky,* 1995; *Comm. of Pennsylvania v. Crawford,* 1996; *Commonwealth v. Kater,* 1983; *Dalrymple v. Brown,* 1997; *Doe v. Maskell,* 1996; *Engstrom v. Engstrom,* 1997; *Hammane v. Humenansky,* 1995; *Hunter v. Brown,* 1996; *John BBB Doe v. Archdiocese of Milwaukee,* 1997; *Kelly et al., v. Marcantonio et al.,* 1996; *Lemmerman v. Fealk,* 1995; *M.E.H. v. L.H.,* 1996, 1997; *People v. Shirley,* 1982; *Ramona v. Ramona,* 1997; *State of New Hampshire v. Hungerford,* 1995, 1997; *State of New Hampshire v. Walters,* 1997; *State of Rhode Island v. Quattrocchi,* 1996; *State v. Atwood,* 1984; *Stokes v. State,* 1989; *S.V. v. R.V.,* 1996; *Travis v. Ziter,* 1996; *Woodroffe v. Hansenclever,* 1995). In the case of "dissociated" abuse memories in MPD patients, there appear to be no well-conducted longitudinal prospective studies showing that severe abuse precedes MPD (whether "dissociated" or not), let alone that such abuse causes MPD.

5. We have the gravest concern about the circumstances and frequency with which MPD diagnoses may be misapplied. There are only two main criteria for DID: two or more personalities that alternately control the patient's behavior, and amnesia. We have observed the diagnosis of MPD being made by clinicians who interpret a patient's varying moods, inconsistent attitudes, and varied likes and dislikes as "alter personalities" where another, more skeptical clinician would not see these behavioral changes as anything out of the ordinary. We have seen reports of clinicians telling patients that any significant forgetting of childhood events (including for events before age 2!) constituted "amnesia," and then using that "symptom" to qualify a patient for an MPD diagnosis. Indeed, supporters of the validity of MPD report that patients often do not manifest any MPD-type behavior when first seen (Kluft, 1984). Most have been treated by various clinicians and for several years under other diagnoses (Coons, Bowman, & Milstein, 1988; Putnam, Guroff, Silberman, Barban, & Post, 1986), with the MPD diagnosis often only made after months of contact with the patient and sometimes made without the symptoms being present (Kluft, 1984).

Under the current state of the art "experts" cannot offer responsible testimony to courts about what conditions favor accurate MPD diagnoses. Much clearer, objective behavioral or physiological criteria are needed. MPD advocates and skeptics alike must generally accept these criteria. Advocates and skeptics also must be able to routinely agree on the presence or absence of symptoms, in individual cases. Until that time, expert testimony supporting MPD diagnoses should not survive an appropriately conducted *Daubert/Kumho* attacks on "standards" grounds.

6. As far as we can tell, MPD/DID is one of the least generally accepted of all *DSM* diagnoses. There are extremely heated debates about the "reality" of MPD/DID, with some defenders (Bliss, 1986; Kluft & Fine, 1993; Putnam, 1989; Ross, 1997) and many skeptics (Aldridge-Morris, 1989; Merskey, 1992; North, Ryall, Ricci & Wetzel, 1993; Ofshe & Watters, 1996; Piper, 1997; Spanos, 1996). We prefer to think of the controversy not in terms of the "reality" of MPD, but rather in terms of its prevalence (near-zero vs. much higher) and its etiology. As a diagnostic label, MPD may conceivably serve as a reasonable label for a certain behavioral syndrome. This syndrome may be rare or more common. It remains to be determined whether it is primarily caused by abuse (Putnam, 1989), by misguided assessment and therapy techniques (Merskey, 1992; Ofshe & Watters, 1996; Piper, 1997), or by media influences (North et al., 1993; Ofshe & Watters, 1996) or is typically feigned or represents a role enactment (Spanos, 1996).

In a recent study by Pope, Oliva, Hudson, Bodkin, and Gruber (1999), U.S. psychiatrists were asked about their attitudes toward MPD. About a third (35%) of respondents favored including MPD in the next *DSM* "without reservations"; 15% thought it should not be included at all; and 43% thought it should be included "only with reservations." In an earlier report, Dell (1988) surveyed professionals treating MPD patients and found that "a total of 52% of the psychiatrists, [and] 80% of the psychologists . . . stated that other professionals had overtly told them that 'there is no such thing as multiple personality [disorder]' " (p. 529). Clearly there is and has been no general acceptance of MPD among general mental health professionals, let alone scientific researchers.

It has been claimed by a few that the relevant scientific community for

Daubert purposes is limited to those who publish research on numerous MPD patients. This superficially plausible requirement conceals several tendentious errors. First, MPD skeptics are highly unlikely to treat many MPD patients, if they view the disorder as created by improper therapy that focuses attention on MPD symptoms. Hence, like most clinicians in general, MPD skeptics are not in a position to see, let alone publish, research on significant numbers of MPD patients. Instead of only listening to clinicians publishing numerous MPD cases, courts should rely on experts in diagnosis and assessment, uses of the *DSM,* and the history of psychiatry to provide guidance regarding the admissibility of MPD.

A second problem with such a restriction is that pro-MPD researchers often publish in "specialty" journals such as the now apparently defunct *Dissociation* (Barach, 1999)—a fact that arguably supports skepticism about the quality of these articles. Third, of the few clinicians who have reported treating dozens to hundreds of MPD patients, a number have been sued or professionally sanctioned for their diagnosis and treatment methods.[1] Restricting the relevant community to such individuals seems unsound. In our view, an appropriate community for deciding on *Daubert* "general acceptance" of MPD would include several different kinds of scientists. It should, of course, include researchers who study MPD patients, but it should also include researchers on disorders frequently present before, after, alongside, or arguably instead of MPD in ostensible MPD patients (e.g., affective disorders, anxiety disorders, personality disorders) as well as diagnostic researchers. It should include developmental psychologists studying children's responses to trauma and experts in human memory (because MPD allegedly involves amnesia). It should include experts in the history of misadventures in psychiatry (because critics allege that MPD is another such misadventure). Finally, it should include psychological and sociological experts in social influence (relevant to judgments about ways MPD could result from suggestion).

How can the MPD concept become generally accepted? It is clearly up to proponents of a controversial theory to win over the scientific community with convincing data. Those proposing a high rate of MPD necessarily suggest that there are up to millions of hitherto hidden cases of MPD now coming to light (Piper, 1997); that such a blatant condition was missed by so many for so long needs a convincing explanation. Those advocating extreme abuse-etiology theories (CIA mind control programming, globe-spanning Satanic cults) must explain how these conspiracies have so successfully and uniformly escaped forensic detection when "ordinary" child molesters, serial killers, the Mafia's crimes, and White House-directed burglaries have not. Pro-MPD theorists must satisfactorily explain why thousands of trauma victims, longitudinally studied over the past several decades, have not been reported to develop MPD (Beitchman et al., 1992; Kendall-Tackett et al., 1993).

Investigators in the fields of child psychology, sociology, history of psychiatry, general clinical psychiatry and clinical psychology, and memory research commonly advocate a simpler alternative explanation for MPD—improper suggestion, chiefly by zealous and poorly trained therapists. The theory that MPD is the result of dissociated memories of horrific abuse is no more reliable than the theory of repression that so many courts have excluded as junk science. We conclude that the

[1]A list is available from the authors on request.

MPD concept, and especially the abuse theory of MPD etiology, does not currently come close to general acceptance under *Daubert/Kumho* analyses.

Summary and Recommendations

The foregoing has implications for courts, expert witnesses, and attorneys. Following *Daubert/Kumho,* federal judges are now on notice by the U.S. Supreme Court that they bear an affirmative duty to actively exclude junk science testimony and thus protect the integrity of the legal process. We hope that state courts quickly follow this overdue analysis, because the admission of unreliable, junk science "expert" testimony is contrary to public policy and endangers the integrity of the legal system. We believe too many citizens have been harmed by inappropriate, unscientific testimony. Proper implementation of *Daubert/Kumho* analyses will go far toward correcting this serious social problem.

We have covered just a few examples of testimonial areas that would, we believe, fail careful *Daubert/Kumho* scrutiny. One cannot generalize to all mental health testimony from these examples. However, given the relatively rigorous requirements of *Daubert,* the recent extension of its reach by *Kumho,* and the limited scientific knowledge base in many areas of clinical psychology and psychiatry, we believe a significant portion of mental health-related social science testimony may have trouble withstanding a well-conducted *Daubert/Kumho* hearing.

In our view, experts have an affirmative ethical duty to refuse to give testimony that would not reasonably be expected to pass *Daubert/Kumho* scrutiny. This is true even if opposing counsel do not challenge its admissibility. In addition, experts have a similar duty to accurately report to judges and juries the kind of research we have summarized. This is true even if opposing counsel does not introduce such research. We say this because professionals are bound by their oath to tell the whole truth; trying to "slip one by" opposing counsel is hardly that. Moreover, unethical testimony (American Psychological Association [APA], 1992; APA Division 41, 1991) inevitably brings the profession into disrepute (Hagen, 1997).

Experts wishing to practice competently in a well-conducted *Daubert/Kumho* hearing will find the new environment a spur to improving their testimony about complex science issues. By contrast, careless experts in *Daubert/Kumho* cross-examinations may reveal culpable technical and ethical errors. It is up to experts to uphold the highest standards of their respective professions, disclose fully and fairly the bases for their opinions, rely to the greatest extent possible on solid scientific findings, explain in understandable terms the uncertainties in their opinions, and be frank about the degree to which their theories and methods meet, or fail to meet, *Daubert* requirements.

Knowledgeable experts can also contribute in another way. Helping educate legal and mental health professionals about scientific methods and about the distinction between good science and junk helps guard the legal system from many forms of bogus expert testimony.

Attorneys may also need to modify traditional practices, where *Daubert/Kumho* hearings are involved. The common practice of "getting up to speed" by a rapid reading of general tomes is likely to prove inadequate in highly complex, science-intensive hearings dealing with multivariate analysis, sophisticated research methodologies, philosophy of science, and detailed facts from decades of research findings (Krauss & Sales, 1999). In the world of *Daubert/Kumho* analysis, a science–law team should be the minimal standard of legal practice, to help ensure that these complexities are properly addressed. We believe attorneys have an affirmative duty to consult with or defer to expert attorneys or scientists in the relevant fields. Improper loss of a *Daubert/Kumho* hearing may yield dire consequences for clients (e.g., false imprisonment of an innocent client, or an innocent child's continued exposure to an abusive environment) and could even lead to a new area of legal malpractice claims. The demand for specialized education and knowledge created by *Daubert, Kumho,* and related decisions is likely to hasten the advent of the multidisciplinary team approach to science-intensive litigation.

References

Albert, S., Fox, H. M., & Kahn, M. W. (1980). Faking psychosis on the Rorschach: Can expert judges detect malingering? *Journal of Personality Assessment, 44,* 115–119.

Aldridge-Morris, R. (1989). *Multiple personality: An exercise in deception.* London: Erlbaum.

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

American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47,* 1597–1611.

American Psychological Association Division 41 and American Psychology-Law Society (1991). Specialty guidelines for forensic psychologists. *Law and Human Behavior, 15,* 655–665.

Barach, P. (1999). President's message. *ISSD News, 17,* 1.

Barrett v. Hyldburg, Super. Ct., Buncombe Co., North Carolina, No. 94-CVS-0793 (January 23, 1996).

Beitchman, J. H., Zucker, K. J., Hood, J. E., daCosta, G. A., Ackman, D., & Cassavia, E. (1992). A review of the long-term effects of child sexual abuse. *Child Abuse and Neglect, 16,* 101–118.

Bersoff, D. N., Glass, D. J., Dodds, L. D., Eckl, L., & Peters, L. M. (1999). *The admissibility of forensic psychological and social science evidence.* Unpublished list of cases.

Blackowiak v. Kemp, 546 N.W. 2d 1, Minn. (April 19, 1996).

Bliss, E. L. (1986). *Multiple personality, allied disorders, and hypnosis.* New York: Oxford University Press.

Borawick v. Shay, 68 F. 3d 597, 2nd Cir., Conn., cert. denied (October 17, 1995)

Bowman, M. (1997). *Individual differences in post traumatic response: Problems with the adversity-distress connection.* Mahwah, NJ: Erlbaum.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Manual for administration and scoring of the MMPI–2.* Minneapolis: University of Minnesota Press.

Carlson v. Humenansky, No. CX-93-7260, 2nd Dist., Ramsey Co., Minn. (December 29, 1995).

Comm. of Pennsylvania v. Crawford, 682 A. 2d 323. Pa. Super. (July 30, 1996).

Commonwealth v. Kater, 447 N.E. 2d 1190 (Mass. 1983).

Coons, P. M., Bowman, E. S., & Milstein, V. (1988). Multiple personality disorder: A clinical investigation of fifty cases. *Journal of Nervous and Mental Disease, 176,* 519–527.

Cronbach, L. J. (1949). Statistical methods applied to Rorschach scores: A review. *Psychological Bulletin, 46,* 393–429.

Dalrymple v. Brown, 1997 WL 499945, Pa., Aug. 25 (1997).

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

Dell, P. (1988). Professional skepticism about multiple personality. *The Journal of Nervous and Mental Disease, 176,* 528–531.

Doe v. Maskell, 679 A. 2d 1087 Md., July 29, 1996, cert. denied 117 S.Ct. 770 (1997).

Engdahl, B., Dikel, T. N., Eberly, R., & Blank, A., Jr. (1997). Post traumatic stress disorder in a community group of former prisoners of war: A normative response to severe trauma. *American Journal of Psychiatry, 154,* 1576–1581.

Engstrom v. Engstrom, No. B098146, Cal. App., 2nd App. Dist., Div. 2, June 18, 1997, unpublished, cert. denied Cal. (September 3, 1997).

Exner, J. E., Jr. (1980). But it's only an *inkblot. Journal of Personality Assessment, 44,* 563–577.

Exner, J. E., Jr. (1993). *The Rorschach: A comprehensive system: Vol. 1. Basic foundations* (3rd ed.). New York: Wiley.

Frye v. United States, 293 F. 1013 (D.C. Cir. 1923).

Galante, R., & Foa, D. (1987). An epidemiological study of psychic trauma and treatment effectiveness for children after a natural disaster. *Journal of the American Academy of Child Psychiatry, 25,* 357–363.

Garb, H. N. (1985). The incremental validity of information used in personality assessment. *Clinical Psychology Review, 4,* 641–655.

Garb, H. N., Florio, C. M., & Grove, W. M. (1998). The validity of the Rorschach and the MMPI. *Psychological Science, 9,* 402–404.

Garb, H. N., Wood, J. M., Nezworski, M. T., Grove, W. M., & Stejskal, W. J. (in press). Towards a resolution of the Rorschach controversy. *Psychological Assessment.*

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

Gianelli, P. C. (1980). The admissibility of novel scientific evidence: *Frye v. United States,* a half-century later. *Columbia Law Review, 1197,* 1223–1224.

Grove, W. M., Zald, D. H., Hallberg, A. M., Lebow, B., Snitz, E., & Nelson, C. (in press). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment.*

Hagen, M. A. (1997). *Whores of the court: The fraud of psychiatric testimony and the rape of American justice.* New York: Harpercollins.

Hammane v. Humenansky, No. C4-94-203, 2nd Dist., Ramsey Co., Minn. (1995).

Hammond, C. (1992, June). *Hypnosis in MPD and ritual abuse.* Paper presented at the Fourth Annual Eastern Regional Conference on Abuse and Multiple Personality, Alexandria, VA.

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

Hunter v. Brown, 1996 WL 57944, Tenn. App., slip copy (February 13, 1996).

Impara, J. C., & Plake, B. S. (Eds.). (1988). *Thirteenth mental measurements yearbook.* Lincoln, NE: Buros Institute.

Jensen, A. R. (1965). A review of the Rorschach. In O. K. Buros (Ed.), *Sixth mental measurement yearbook* (pp. 501–509) Highland Park, NH: Gryphon.

John BBB Doe v. Archdiocese of Milwaukee, 565 N.W. 2d 94, Wisc., June 27, 1997.

Kelly et al. v. Marcantonio et al., 678 A. 2d 873, R.I. (July 11, 1996).

Kendall-Tackett, K. A., Williams, L. M., & Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. *Psychological Bulletin, 113,* 164–180.

Kluft, R. P. (1984). Introduction to multiple personality disorder. *Psychiatric Annals, 14,* 19–24.

Kluft, R. P., & Fine, C. G. (1993). *Clinical perspectives on multiple personality disorder.* Washington, DC: American Psychiatric Press.

Krauss, D. A., & Sales, B. D. (1999). The problem of "helpfulness" in applying *Daubert* to expert testimony: Child custody determinations in family law as an exemplar. *Psychology, Public Policy, and Law, 5,* 78–99.

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

Lemmerman v. Fealk, 534 N.W. 2d 695, Mich. (July 5, 1995).

Lipton, J. P. (1999). The use and acceptance of social science evidence in business litigation after *Daubert. Psychology, Public Policy, and Law, 5,* 59–77.

Little, K. B., & Schneidman, E. S. (1959). Congruencies among interpretations of psychological test and anamnestic data. *Psychological Monographs, 73*(6, Whole No. 476).

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.

McFarlane, A. C. (1987). Life events and psychiatric disorder: The role of a natural disaster. *British Journal of Psychiatry, 151,* 362–367.

McFarlane, A. C. (1988a). The aetiology of post-traumatic stress disorders following a natural disaster. *British Journal of Psychiatry, 152,* 116–121.

McFarlane, A. C. (1988b). The longitudinal course of posttraumatic morbidity: The range of outcomes and their predictors. *Journal of Nervous & Mental Disease, 176,* 30–39.

McFarlane, A. C. (1988c). The phenomenology of posttraumatic stress disorders following a natural disaster. *Journal of Nervous & Mental Disease, 176,* 22–29.

McFarlane, A. C. (1988d). Relationship between psychiatric impairment and a natural disaster: The role of distress. *Psychological Medicine, 18,* 129–139.

McHugh, P. R. (1998). Testimony given in *Daubert* hearing, *State of Rhode Island v. John Quattrocchi.*

M.E.H. v. L.H., 669 N.E. 2d 1228, 218 Ill. Dec. 702, Ill. App. 2nd Dist., affirmed by Ill. Sup. Ct. (August 28, 1996).

M.E.H. v. L.H., 1997 WL 562001, Ill., slip copy (September 4, 1997).

Meloy, J. R. (1991). Rorschach testimony. *Journal of Psychiatry and Law, 8,* 221–235.

Merskey, H. (1992). The manufacture of personalities: The production of multiple personality disorder. *British Journal of Psychiatry, 160,* 327–340.

National Prohibition Act, 41 Stat. 305 (1919) (enforcing the Eighteenth Amendment to the U.S. Constitution).

Nezworski, M. T., & Wood, J. M. (1995). Narcissism in the Comprehensive System for the Rorschach. *Clinical Psychology: Science and Practice, 2,* 179–199.

Nolen-Hoeksema, S., & Morrow, J. (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: The 1989 Loma Prieta earthquake. *Journal of Personality and Social Psychology, 61,* 115–121.

North, C. S., Ryall, J.-E. M., Ricci, D. A., & Wetzel, R. D. (1993). *Multiple personalities: Psychiatric classification and media influence.* New York: Oxford University Press.

Ofshe, R., & Watters, E. (1996). *Making monsters: False memories, psychotherapy, and sexual hysteria.* Berkeley: University of California Press.

Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Consulting Psychology, 29,* 261–265.

Parker, K. C. H., Hanson, R. K., & Hunsley, J. (1988). MMPI, Rorschach, and WAIS: A meta-analytic comparison of reliability, stability, and validity. *Psychological Bulletin, 103,* 367–373.

People v. Shirley, 31 Cal. 3d 18 (1982).

Piper, A., Jr. (1997). *Hoax and reality: The bizarre world of multiple personality disorder.* Northvale, NJ: Jason Aronson.

Pope, H. G., Jr., & Hudson, J. I. (1995). Can individuals "repress" memories of childhood sexual abuse? An examination of the evidence. *Psychiatric Annals, 25,* 715–719.

Pope, H. G., Jr., Hudson, J. I., Bodkin, J. A., & Oliva, P. S. (1998). Questionable validity of "dissociative amnesia" in trauma victims: Evidence from prospective studies. *British Journal of Psychiatry, 172,* 210–215.

Pope, H. G., Jr., Oliva, P. S., & Hudson, J. I. (1999). The scientific status of research on repressed memories. In D. L. Faigman, D. H. Kaye, M. J. Saks, & J. Sanders (Eds.), *Modern scientific evidence: The law and science of expert testimony* (Vol. 1, pp. 115–155). St. Paul, MN: West Group.

Pope, H. G., Jr., Oliva, P. S., Hudson, J. I., Bodkin, J. A., & Gruber, A. J. (1999). Attitudes toward DSM–IV dissociative disorders diagnoses among board-certified American psychiatrists. *American Journal of Psychiatry, 156,* 321–323.

Popper, K. R. (1959). *The logic of scientific discovery.* New York: Basic Books.

Putnam, F. W. (1989). *Diagnosis and treatment of multiple personality disorder.* New York: Guilford Press.

Putnam, F. W., Guroff, J. J., Silberman, E. K., Barban, L., & Post, R. M. (1986). The clinical phenomenology of multiple personality disorder: Review of 100 recent cases. *Journal of Clinical Psychiatry, 47,* 285–293.

Ramona v. Ramona, 66 Cal. Rptr. 2d 766, Ca. App. (August 19, 1997).

Ross, C. A. (1997). *Dissociative identity disorder: Diagnosis, clinical features, and treatment of multiple personality* (2nd ed.). New York: Wiley.

Schopp, R., Scalora, M. J., & Pearce, M. (1999). Expert testimony and professional judgment: Psychological expertise and commitment as a sexual predator after *Hendricks. Psychology, Public Policy, and Law, 5,* 120–174.

Spanos, N. P. (1996). *Multiple identities and false memories: A socio-cognitive perspective.* Washington, DC: American Psychological Association.

Spitzer, R. L. (1983). Psychiatric diagnosis: Are clinicians still necessary? *Comprehensive Psychiatry, 24,* 399–411.

State of New Hampshire v. Hungerford, 1995 WL 378571, N.H. Super. Ct. (May 23, 1995).

State of New Hampshire v. Hungerford, 1997 WL 358620, N.H. (July 1, 1997).

State of New Hampshire v. Walters, 1997 WL 937024, N.H. (August 6, 1997).

State of Rhode Island v. Quattrocchi, 681 A. 2d 879, R.I. (1996).

State v. Atwood, 479 A. 2d 258 (Conn. 1984).

Stokes v. State, 548 So. 2d 1988 (Fla. 1989).

S. V. v. R. V., 933 S.W. 2d 1, 39 Tex. Supp. J. 386, Tex. (March 14, 1996).

Tennant, C. (1983). Life events and psychological morbidity: The evidence from prospective studies. *Psychological Medicine, 13,* 483–486.

Tennant, C., Bebbington, P., & Hurry, J. (1981). The role of life events in depressive illness: Is there a substantial causal relation? *Psychological Medicine, 11,* 379–389.

Travis v. Ziter, 681 So. 2d 1348, Ala. (July 12, 1996).

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

Wood, J. M., Lilienfeld, S. O., Garb, H. N., & Nezworski, M. T. (in press). *The Rorschach test in clinical diagnosis: A critical review, with a backward look at Garfield (1947).* Journal of Clinical Psychology.

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

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

Woodroffe v. Hansenclever, 540 N.W. 2d 45, Iowa, Nov. 22, 1995.

Received June 8, 1999
Revision received August 20, 1999
Accepted August 28, 1999 ■