

Protecting Children From Incompetent Forensic Evaluations and Expert Testimony

Mary Johanna McCurley*
Kathryn J. Murphy**
Jonathan W. Gould***

I. Introduction

Mental health professionals are frequently appointed by courts to become involved in custody cases in the role of child custody evaluator. This role requires that the mental health professional assess the fit between a minor child's emerging developmental and socioemotional needs and the parents' comparative ability to meet those needs. Following that assessment, the mental health professional is expected to tender recommendations to the court regarding the extent to which various parenting plans will further the child's best psychological interests.

A. *Influence of the Evaluator*

The recommendations contained in child custody evaluations ("CCEs") exert considerable influence on the course of ongoing custody litigation. Many courts accord significant weight to the opinions of child custody evaluators, often accepting the evaluator's recommendations without challenge.¹ An evaluator's recommendations can also precipitate case settlement or material concessions once both parties become aware of the evaluator's findings. Given the import of CCEs, it is imperative that these

* Partner, McCurley, Orsinger, McCurley, Nelson & Downing, L.L.P., Dallas, Texas.

** Partner, Koons, Fuller, Vanden Eykel, and Robertson, P.C., Dallas, Texas.

*** Forensic and Clinical Psychologist, Private Practice, St. Paul, Minnesota.

¹ THE SCIENTIFIC BASIS OF CHILD CUSTODY DECISIONS (Robert Galatzer-Levy & Louis Kraus, eds., 1999); James N. Bow & Francella A. Quinnell, *A Critical Review of Child Custody Evaluation Reports*, 40 FAM. CT. REV. 164 (2002).

evaluations should be conducted with due regard for scientific methods, extant behavioral science research, ethical standards and professional practice guidelines. This responsibility is further heightened by the need to protect the well-being of children of divorcing parents² and to avoid inadvertently harming families at a time of enormous stress and conflict.³

B. *Criticism of Evaluations Falling Below APA Standards*

Unfortunately, CCEs frequently fall below recommended practice methods promulgated by the American Psychological Association (“APA”). Commentators have criticized the quality, reliability, and utility of CCEs by noting the lack of scientific methodology, empirical grounding, and psycholegal relevance common among these reports.⁴

This article discusses the requirement that expert testimony regarding parenting competency and comparative custodial suitability must meet legal standards of admissibility. It builds on a prior article coauthored by Dr. Jonathan Gould.⁵

² ROBERT H. WOODY, *CHILD CUSTODY: PRACTICE STANDARDS, ETHICAL ISSUES, AND LEGAL SAFEGUARDS FOR MENTAL HEALTH PROFESSIONALS* (2000).

³ Marc J. Ackerman & Melissa C. Ackerman, *Child Custody Evaluation Practices: A 1996 Survey of Psychologists*, 30 *FAM. L.Q.* 565 (1996).

⁴ THE SCIENTIFIC BASIS OF CHILD CUSTODY DECISIONS, *supra* note 1; GARY B. MELTON, JOHN PETRILA, NORMAN G. POYTHRESS, & CHRISTOPHER SLOBOGIN, *PSYCHOLOGICAL EVALUATIONS FOR THE COURTS: A HANDBOOK FOR MENTAL HEALTH PROFESSIONALS AND LAWYERS* (2nd ed. 1997); Daniel A. Krauss & Bruce D. Sales, *Legal Standards, Expertise, and Experts in the Resolution of Contested Child Custody Cases*, 6 *PSYCHOL., PUB. POL’Y & L.* 843 (2000); Timothy M. Tippins & Jeffrey P. Wittman, *Empirical and Ethical Problems with Custody Recommendations: A Call for Clinical Humility and Judicial Vigilance*, 43 *FAM CT. REV.* 193 (April 2005); Ira D. Turkat, *Questioning the Mental Health Expert’s Custody Report*, 7 *AM. J. FAM. L.* 175 (1993).

⁵ The authors wish to thank Dana Royce Baerger, Robert Galatzer-Levy, and Sandra G. Nye for allowing us to liberally use their prior article: Dana Royce Baerger, Robert Galatzer-Levy, Jonathan W. Gould, & Sandra G. Nye, *Methodology for Reviewing the Reliability and Relevance of Child Custody Evaluations*, 18 *J. AM. ACAD. MATRIM. LAW.* 35 (2002).

C. Differences Between Therapeutic and Forensic Mental Health Assessment

Some child custody evaluators do not adequately understand the distinctions between a therapeutic and a forensic role. Furthermore, much of the testimony offered by child custody evaluators is based upon clinical impressions uninformed by empirical research, yet presented as empirical science. These two points merit further discussion. First, as the field of forensic mental health has grown to constitute a distinct subspecialty, commentators have noted the increasing need for practitioners to obtain specialized and advanced training.⁶ The reason for this is that the field requires appreciably distinct competencies and skills than does the treatment of patients. Moreover, differences between forensic and therapeutic services have now been codified in ethical guidelines⁷ and clarified in the behavioral science literature.⁸

II. Expert Testimony in Child Custody Evaluations — *Daubert/Frye*

A. *Frye v. United States*

Expert testimony has long played a substantial role in the trial of a child custody case. Whether in the form of a social study or a psychiatric evaluation after a battery of tests, testimony based on the social sciences has become the norm. Expert testimony regarding parenting competency and comparative custodial suitability must meet legal standards of admissibility.

In 1923, the U.S. Supreme Court in *Frye v. United States*⁹ issued the primary determinative test for the admissibility of “novel” scientific evidence. For seventy years, the general stan-

⁶ Randy K. Otto & Kirk Heilbrun, *The Practice of Forensic Psychology: A Look Toward the Future in Light of the Past*, 57 AM. PSYCHOLOGIST 5 (2002).

⁷ AMERICAN ACADEMY OF PSYCHIATRY AND THE LAW, ETHICAL GUIDELINES FOR THE PRACTICE OF FORENSIC PSYCHIATRY (1995); Committee on Ethical Guidelines for Forensic Psychologists, *Specialty Guidelines for Forensic Psychologists*, 15 LAW & HUM BEHAV. 655 (1991).

⁸ Lyn R. Greenberg & Jonathan W. Gould, *The Treating Expert: A Hybrid Role with Firm Boundaries*, 32(5) PROF. PSYCHOL.: RES. & PRAC. 469 (2001).

⁹ 293 F. 1013 (D.C. Cir. 1923).

dard for admissibility of expert testimony in most jurisdictions was the *Frye* test, which permits admission of evidence when the foundational scientific principle has achieved “general acceptance” in the scientific field.¹⁰

In *Frye*, defense counsel sought to introduce expert testimony explaining the results of a test, similar to today’s polygraph test. In upholding the trial court’s refusal to admit the expert evidence, the court of appeals stated that while courts will go a long way in admitting expert testimony, “the thing from which the [expert testimony is deduced must be] sufficiently established to have gained general acceptance in the particular field in which it belongs.”¹¹ Even though the Court cited no authority to support this statement, the “general acceptance” standard became the cornerstone for the admissibility of novel scientific evidence in the federal courts and most of the state courts for many years.

The *Frye* standard has generated and continues to provoke great debate and commentary.¹² Proponents advocate that the *Frye* “general acceptance” standard is the proper test for admissibility because: its conservative nature ensures the reliability of novel scientific evidence; it better promotes uniformity of decisions; scientific evidence tends to have a substantial impact on a jury; without a standard test to determine admissibility, trials could become mini-trials concerning the scientific evidence, thus distracting the jury from the merits of the case to be tried; and there will be a reserve of experts who may be called upon to express an opinion regarding the validity of the evidence.¹³

After adoption of the Federal Rules of Evidence in 1975, the issue of the appropriate standard for admissibility of scientific evidence became confusing. The federal circuits and state courts were split on what the standard should be. The Rules had been in existence for eighteen years by the time the U.S. Supreme Court announced a novel standard for the admissibility of scientific evidence. Immediately before the end of the 1993 term, the Supreme Court unanimously rejected the *Frye* test, stating that it

¹⁰ *Id.* at 1014.

¹¹ *Id.*

¹² *E.g.*, Mike McCurley & C. Andrew Ten Eyck, *Daubert and the Admissibility of Expert Testimony in Child Custody Disputes*, 19 No. 2 *MATRIM. STRATEGIST* 1 (2001).

¹³ *See, e.g.*, *Reed v. State*, 391 A.2d 364, 369-372 (Md. 1978).

had been superseded by the Federal Rules of Evidence. Further, a majority of the Court suggested new standards for the admissibility of scientific evidence in the pivotal case of *Daubert v. Merrill Dow Pharmaceuticals, Inc.*¹⁴

B. *Daubert v. Merrill Dow Pharmaceuticals, Inc.*

In *Daubert*, the Court reinterpreted Rule 702 of the Federal Rules of Evidence and established criteria for the admissibility of scientific expert testimony. Amended Rule 702 reads:

If *scientific*, technical, or other specialized *knowledge* will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.¹⁵

Although numerous states have followed *Daubert*,¹⁶ many have not. States that have rejected *Daubert* in favor of *Frye*, or an alternative standard, have expressed concerns about the reliability and relevance of proffered expert testimony.¹⁷

¹⁴ 509 U.S. 579 (1993).

¹⁵ FED. R. EVID. 702 (emphasis added).

¹⁶ As of 2004, all but 15 states have adopted *Daubert* or some variation of *Daubert* to examine whether an expert's testimony is admissible as scientific evidence. Alice B. Lustre, Annotation, *Post-Daubert Standards for Admissibility of Scientific and Other Expert Evidence in State Courts*, 90 A.L.R.5th 453, 520-37 (2004). Some states, such as New York, apply the *Daubert* examination in limited circumstances. The largest state to continue to adhere to the *Frye* approach is California.

¹⁷ Because not all states have adopted the *Daubert* standard for admissibility of expert testimony, and because some states have retained the *Frye* general acceptance criterion, or some modification thereof, or are using a combination of *Frye* and *Daubert*, it is essential to attend to the evidentiary standards of the jurisdiction in which the testimony is given. Generally, even in locales where *Frye* and similar standards continue in force, custody evaluators are increasingly expected to base their opinions on scientifically sound methods. At a minimum, evaluators should clearly state the methodological basis for their opinions. See Daniel W. Schuman, *The Role of Mental Health Experts in Custody Decisions: Science, Psychological Tests, and Clinical Judgment*, 36 FAM. L.Q. 135 (2002).

In jurisdictions that continue to use the *Frye* rule, the proponent of the expert testimony must establish the qualifications of the expert, as well as that

According to *Daubert*, a “scientific expert” is an expert who relies upon the application of scientific principles, rather than upon skill or experience, in forming his or her opinions. Moreover, according to the Supreme Court’s later pronouncement in *Kumho Tire Co., Ltd. v. Carmichael*,¹⁸ *Daubert* applies to all expert testimony, not just to scientific testimony. Thus, in *Daubert* jurisdictions, the distinction between scientific and nonscientific expert testimony is not significant in establishing admissibility; the same criteria apply to both. Moreover, the admissibility of expert testimony based upon personal observation and clinical experience is subject to judicial scrutiny regarding its reliability. Expert testimony that goes to the causation of a condition is also subject to scrutiny for reliability.¹⁹

The *Daubert* court made clear that when it used the term “reliability,” it was referring both to scientific reliability and scientific validity.²⁰ Essentially, scientific reliability refers to consistency. In the realm of psychological testing, scientific reliability refers to the consistency of test scores. For example, assume that a psychologist wishes to construct a new measure of stable per-

the espoused theory has gained general acceptance in the expert’s scientific field. When the expert testimony being offered is not scientific, but rather is based upon observation and experience, *Frye* does not apply and no *Frye* hearing is required. However, the court may still conduct a hearing on the reliability of the testimony being offered. *Berry v. City of Detroit*, 25 F.3d 1342, 1349-50 (6th Cir. 1994). In a *Frye* jurisdiction, counsel should be prepared to lay the following foundation for admissibility of expert testimony: (1) the evidence that is being proffered; (2) the fact that the proffered testimony will assist the trier of fact in understanding the evidence or determining facts in issue (and the way in which the proffered testimony will so assist); (3) whether the proffered testimony constitutes ‘science;’ and (4) the qualifications of the expert (including knowledge, skill, training, experience, and education) to testify and render an opinion. *See, e.g., In re Jawad*, 759 N.E.2d 1002 (Ill. App. Ct. 2001). Whether or not *Daubert* has replaced *Frye* in a given jurisdiction, the concepts established by *Daubert* seem to be finding their way into *Frye* jurisdictions. *See Harris v. Cropmate Co.*, 706 N.E.2d 55 (Ill. App. Ct. 1999), for a comparison of *Daubert* and *Frye* that describes the evolution of *Frye* into a “*Frye* plus reliability” standard for the introduction of novel scientific evidence.

¹⁸ 526 U.S. 137 (1999).

¹⁹ *See, e.g., Commonwealth v. Lanigan*, 641 N.E.2d 1342, 1348 (Mass. 1994)(applying both the *Daubert* and *Kumho Tire Co.* criteria); *see also In Re Canovan’s Case*, 733 N.E.2d 1042, 1047-51 (Mass. 2000)(containing a summary of the standards for admission of expert testimony).

²⁰ 509 U.S. at 590.

sonality traits, with responses summed on a scale of zero to one hundred. If a test subject takes the test four separate times, and each time receives a widely divergent score (despite the fact that his or her mood has remained stable), then we can say that the test is not reliable, since the test subject did not receive scores that were consistent over time and the construct being measured was assumed to be consistent over time. Another way to think about reliability is as the answer to the question, “Are we measuring something (e.g., a genuine trait, behavior, factor, or phenomenon) consistently?”

Scientific validity refers to accuracy and utility (mental health professionals use the term “validity” to refer to what legal professionals call “reliability”). In the realm of psychological testing, scientific validity refers to the extent to which the test measures what it purports to measure. In the example above, once the psychologist has established that the new measure of personality traits is reliable, he or she will want to know whether it actually measures a known personality trait—as opposed to, say, mood variability that often changes over time. If that is the case, the measure will not be able to discriminate between test subjects who have a specific personality trait and the test subjects who are displaying mood variability (e.g., the measure will lack “discriminant validity”). Another way to think about validity is as the answer to the question, “Are we measuring what we think we’re measuring?” As the above example makes clear, a test must be scientifically reliable to be scientifically valid, although the reverse is not true. In other words, reliability is a part of what makes a test valid, since if a test cannot consistently measure some factor, then it is highly unlikely that the test can impart any useful information.

A proper *Daubert* challenge may include not only an attack on the entire methodology used by the testifying professional, but specific challenges as to the reliability and validity of each psychological instrument administered. Attorneys involved in custody cases, therefore, must have a general working knowledge of how such tests meet or fail to meet a serious analysis of their reliability and validity.

Daubert has had a significant impact on the admissibility of behavioral and psychological evidence offered by expert psychol-

ogist witnesses, mainly in child abuse and molestation cases.²¹ Experts who routinely testify in divorce and custody cases sometimes report that their efforts to testify regarding novel theories have been thwarted by trial courts based on the restrictive guidelines of *Daubert*.

To defend against a claim that the psychologist's testimony is inadmissible, the psychologist should make every effort to follow proper procedures for selecting, administering, scoring, and interpreting tests, while taking care to follow all applicable ethical standards and professional practice guidelines for his or her field. In custody cases, the examination of the child and the family as a whole needs to be undertaken with extreme care. The psychologist or psychiatrist needs to understand what his or her role is in the case and make sure that he or she stays within those parameters.

To determine whether testimony about scientific knowledge will assist the trier of fact in assessing a controverted issue, *Daubert* requires that the judge ask two questions: (1) whether the reasoning or methodology underlying the testimony is scientifically valid, and (2) whether that reasoning or methodology can be properly applied to the facts in issue.²²

C. *Federal Rule of Evidence 702*

For a witness to qualify to testify as an expert under Rule 702, the following steps must be taken. First, the expert must be *qualified* to give an opinion.²³ Second, the opinion must be considered *reliable*.²⁴ Third, the testimony must be *relevant* and assist the trier of fact in coming to a decision in the case; in other words, the expert must convey specialized knowledge beyond that of a layperson.²⁵

²¹ See, e.g., *State v. Foret*, 628 So.2d 1116, 1123 (La. 1993)(refusing to accept testimony from a child psychologist regarding the child sexual abuse accommodation syndrome (CSAAS), because the absence of methods to distinguish valid CSAAS claims from manufactured claims meant that testimony based on the syndrome did not survive the *Daubert* test for scientific validity).

²² *Daubert*, 509 U.S. at 592-93.

²³ *Id.* at 592 n.10.

²⁴ *Id.* at 590 n.9.

²⁵ *Id.* at 591-92.

1. *The Expert Must Be Qualified*

The testifying expert should be prepared to testify that he or she possesses the qualifications (education, experience, etc) and knowledge necessary to relate the information to the fact finder. Rule 702 requires that experts be qualified “by knowledge, skill, experience, training, or education” and that their testimony assist the trier of fact in determining a fact issue.²⁶ Whether an expert is qualified is a preliminary question decided by the trial court.²⁷

The party offering the expert must show that the expert’s knowledge, experience, skill, or training or education renders the expert qualified to give an opinion regarding the *specific issue* before the court.²⁸ For example, not every doctor is qualified to testify as an expert on every medical question.²⁹ The party offering the expert’s testimony bears the burden to prove that the witness is qualified under Rule 702. The offering party must demonstrate that the witness possesses special knowledge regarding the very matter on which he proposes to give an opinion. Not every psychologist is qualified to testify on all psychological issues.

If the opinion relates to the standard of care within a licensed profession, the expert will generally be required to be licensed in that same profession. The expert should be familiar with as much of the available literature on the field as possible and be prepared to respond to inquiries on that literature, particularly any publications that are critical of the expert’s own theories or practical application of those theories.

²⁶ FED. R. EVID. 702.

²⁷ FED. R. EVID. 104.

²⁸ Daubert, 509 U.S. at 597 (stating that “an expert’s testimony” [must be] “relevant to the task at hand.”); FED. R. EVID. 401. In proffering evidence as to the qualifications of one’s expert witness, one should elicit from the witness the following types of testimony: the profession or occupation of the witness; the length of time in the profession or occupation; the educational background, including degrees obtained, of the witness; the professional training in which the witness has participated; licensure of the witness; membership of the witness in professional associations; research, publications, writing, or articles by the witness; past experience as an expert witness; and, academic positions if any.

²⁹ See *Christophersen v. Allied-Signal Corp.*, 939 F.2d 1106, 1112-13 (5th Cir. 1991), *cert. denied*, 503 U.S. 912 (1992), in which the Fifth Circuit Court of Appeals states that one possessing a medical degree is not qualified to give an opinion on every medical question.

2. *The Expert's Opinion Must Be Reliable*

Again, it is important to recognize that while courts use the terms “reliability” and “validity” interchangeably, mental health professionals use the terms in distinctly different ways.³⁰ Social scientists consider “reliability” to mean the technical accuracy of a test, while the “validity” of the test is based on whether the test leads to a legitimate conclusion.

To be reliable, the underlying scientific technique or principle must be grounded in the methods and procedure of science. To show reliability, the party who is offering the opinion should address the “non-exclusive” list of factors in *Daubert*,³¹ as well as additional factors addressed in each state’s case law.

a. *Testing*³²

Whether the theory or technique in question can be (and has been) tested (referred to as the “falsifiability” of a theory). This factor recognizes that testing is central to scientific methodology.³³ If a theory or technique has been or can be tested, the trial judge can better determine “whether a theory or technique is scientific knowledge that will assist the trier of fact.”

b. *Peer review*

Whether the theory or technique has been or could be subjected to peer review or publication is an important factor. Publication and other peer review is a significant indicia of the reliability of scientific evidence when the expert’s testimony is in an area in which peer review or publication would not be uncommon. Publication in a reputable, established, scientific journal and other forms of peer review increase the likelihood that substantive flaws in methodology will be detected. Although peer review and publication are indicators of evidentiary reliability,

³⁰ See *Daubert*, 509 U.S. at 590 n.9.

³¹ *Id.* at 592-94.

³² *Id.* at 592 (discussing testing).

³³ “If theories are non-falsifiable, they are unscientific. Hypotheses that avoid testability . . . are not scientific. Likewise, theories that purport to explain everything, theories that are unconditional and admit no negative evidence and theories that are vague or otherwise self-protected are not falsifiable.” Nancy Levit, *Listening to Tribal Legends: An Essay on Law and the Scientific Method*, 58 *FORDHAM L. REV.* 263, 271 (1989).

the *Daubert* Court emphasized that “publication (or lack thereof) in a peer-reviewed journal” is not dispositive of the question of scientific validity.³⁴

c. *Error rates*

The theory or technique’s “known or potential error rate.” In this context, “error rate” refers to the probability that the application of a particular technical procedure or theory can lead to a mistake in classification of an object, event or person. *Daubert* says that trial courts should survey studies of the error rates of the specific technique, as well as the standards controlling the technique’s operation.³⁵

d. *General acceptance*

Whether the theory or technique has been generally accepted as valid by the relevant scientific community. “General acceptance” continues to be a significant persuasive factor in determining whether particular evidence is admissible. It is not, however, the dispositive requirement.³⁶

3. *The Expert’s Opinion Is Relevant*

The expert’s opinion must be relevant. To be relevant, the evidence must have probative value and must have consequence to some issue in the trial. Relevancy is the threshold question to the admissibility of evidence regardless of whether it is the testimony of a fact witness or an expert witness.³⁷ The evidence must be sufficiently tied to the facts of the case that it will aid the jury in resolving a factual dispute. Evidence that has no relationship to any of the issues in the case is irrelevant, and therefore inadmissible under Rules 401, 402, and 702.

Relevance refers to the extent that the gathered data bear upon the issue before the court. In the contest of a CCE, in which the issue before the court is one of comparative custodial suitability, relevant data will include information that relates to the litigants’ parenting strengths and deficits, the child’s relationship with each parent, and the quality of fit between the child’s

³⁴ *Daubert*, 509 U.S. at 594. See also *id.* at 593 (discussing peer review).

³⁵ *Id.* (discussing error rates).

³⁶ *Id.* (discussing general acceptance).

³⁷ *Id.* at 587.

needs and the parents' respective care giving competencies. For example, data relevant to a CCE will include information regarding how each parent supervises, disciplines, supports, nurtures and instructs the child.

Relevance also refers to the extent to which a data-gathering technique is valid (e.g., useful) for the purpose for which it is used. A psychological test can be reliable and valid, and still lack relevance in the context of a CCE because it does not measure, either directly or indirectly, relevant factors associated with psychological factors of concern to the court. For example, with the exception of serious cognitive impairment that would make adequate childcare difficult, no empirically established relationship exists between intellectual functioning and parenting capacity.³⁸ Since no methods are available to discern how a parent's intellectual functioning as identified through an IQ score might impact (if at all) on his or her parenting capabilities, it is unlikely that tests of intellectual functioning will be valid (e.g., useful) for the purpose of determining parenting competencies. In this sense, tests of intellectual functioning are not relevant to the legal issue of comparative custodial suitability.

Reliability and relevance can be illustrated by the multi-trait/multi-method model of assessment. Forensic assessment is predicated upon the idea of convergent validity, or the idea that particular issues should be investigated from a variety of viewpoints and with a variety of methods. Addressing the same issue through a number of different data sources will likely increase the reliability of the information gathered, since the evaluator can then look for consistent trends across the data. Hence, a competently conducted forensic evaluation utilizes multiple sources of information to assess multiple aspects of a situation; this is referred to as the multi-trait/multi-method model of assessment. This model of obtaining convergent data from multiple sources for a CCE has achieved increasing professional consensus over the past ten years, and has been described as the model that best serves the evidentiary needs of the court.³⁹ Further-

³⁸ Thomas Grisso, *The Economic and Scientific Future of Forensic Psychological Assessment*, 42 AM. PSYCHOLOGIST 831 (1987).

³⁹ Ackerman & Ackerman, *supra* note 3, at 567; Jon K. Amundson, Roshni Daya, & Eamon Gill, *A Minimalist Approach to Child Custody Evaluations*, 18 AM. J. FORENSIC PSYCHOL. 63 (2000); James N. Bow & Francella A.

more, this model is included in almost all recommended ethical standards and professional practice guidelines for conducting forensic evaluations, including the American Psychological Association,⁴⁰ the Association of Family and Conciliation Courts,⁴¹ and the Specialty Guidelines for Forensic Psychologists.⁴²

Hence, the objective of a CCE is to assess functional parenting competencies in a reliable and relevant manner. Currently, there exist a number of different CCE models;⁴³ no requisite set of procedures or tests has been defined. However, a consensus is emerging in the behavioral science literature regarding the manner in which CCEs should be conducted and the procedures that are most likely to ensure the assessment's reliability and relevance. Greenberg and Gould⁴⁴ have proposed a five-part methodological framework that synthesizes the legal and behavioral science literature,⁴⁵ empirical research,⁴⁶ ethical guidelines,⁴⁷ and

Quinnell, *Psychologists' Current Practices and Procedures in Child Custody Evaluations: Five Years After the American Psychological Association Guidelines*, 32 PROF. PSYCHOL.: RES. & PRAC. 261 (2001).

⁴⁰ American Psychological Association, *Ethical Principles of Psychologists and Code of Conduct*, 57 AM. PSYCHOLOGIST 1060 (2002).

⁴¹ Association of Family and Conciliation Courts, *Model Standards of Practice for Child Custody Evaluations*, 32 FAM. & CONCILIATION CTS. REV. 504 (1994).

⁴² Committee on Ethical Guidelines for Forensic Psychologists, *supra* note 7.

⁴³ See, e.g., MARC J. ACKERMAN, *CLINICIANS GUIDE TO CHILD CUSTODY EVALUATIONS* (1995); PHILIP M. STAHL, *CONDUCTING CHILD CUSTODY EVALUATIONS: A COMPREHENSIVE GUIDE* (1994).

⁴⁴ Greenberg & Gould, *supra* note 8, at 471-73.

⁴⁵ See generally, Grisso, *supra* note 38; Melton, et al, *supra* note 4 ; see also RICHARD ROGERS, *DIAGNOSTIC AND STRUCTURED INTERVIEWING: A HANDBOOK FOR PSYCHOLOGISTS* (1995); BENJAMIN M. SCHUTZ, ELLEN B. DIXON, JOANNE C. LINDENBERGER, & NEIL J. RUTHER, *SOLOMON'S SWORD: A PRACTICAL GUIDE TO CONDUCTING CHILD CUSTODY EVALUATIONS* (1989); David L. Faigman, *The Evidentiary Status of Social Science under Daubert: Is it 'Scientific,' 'Technical,' or 'Other' Knowledge?*, 1 PSYCHOL., PUB. POL'Y, & L. 960 (1995); David L. Faigman, *Struggling to Stop the Flood of Unreliable Expert Testimony*, 76 MINN. L. REV. 877 (1992); Jane Goodman-Delahunty, *Forensic Psychological Expertise in the Wake of Daubert*, 21 LAW & HUM. BEHAV. 121 (1997); Kirk Heilbrun, *The Role of Psychological Testing in Forensic Assessment*, 16 LAW & HUM BEHAV. 257 (1992).

⁴⁶ See, e.g., Marc J. Ackerman & Melissa C. Ackerman, *Custody Evaluation Practices: A Survey of Experienced Professionals (Revisited)*, 28 PROF. PSYCHOL.: RES. & PRAC. 137 (1997); Peter Ash & Melvin J. Guyer, *Biased*

model standards of practice⁴⁸ regarding CCEs. This framework for CCEs incorporates five core data-gathering components:

- (1) a definition of the scope of the evaluation;
- (2) the use of forensic interview techniques;
- (3) psychological testing with objective and self-report measures;
- (4) direct behavioral observations of parent-child interactions; and
- (5) interviews with collateral sources and review of relevant records.

Consistent with *Daubert*, this approach to CCEs assesses parental competencies within the parameters of reliability and relevance, and thus constitutes a genuinely scientific methodology. Moreover, this approach can also be used as an organizational structure with which to evaluate the methodological strengths and shortcomings of other CCEs.⁴⁹

III. Five Core Components for Data Gathering

Following is an explication of Gould's five-part methodology for an analysis of the reliability and relevance of CCEs.

A. Define the Scope of the Evaluation

When a forensic mental health professional begins a CCE, he or she should identify the specific questions to be investigated prior to the initiation of the evaluation.⁵⁰ In this way, the evalu-

Reporting by Parents Undergoing Child Custody Evaluations, 30 J. AM. ACAD. CHILD & ADOLESCENT PSYCHIATRY 835 (1991); Kay Bathurst, Allen W. Gottfried, & Adele E. Gottfried, *Normative Data for the MMPI-2 in Child Custody Litigation*, 9 PSYCHOL. ASSESSMENT 205 (1997); Chery Hysjulien, Barbara Wood, & G. Andrew H. Benjamin, *Child Custody Evaluations: A Review of Methods Used in Litigation and Alternative Dispute Resolution*, 32 FAM. & CONCILIATION CTS. REV. 466 (1994).

⁴⁷ See generally American Psychological Association, *supra* note 40; Committee on Ethical Guidelines for Forensic Psychologists, *supra* note 7.

⁴⁸ See generally Association of Family and Conciliation Courts, *supra* note 45.

⁴⁹ Jonathan W. Gould & Lisa C. Bell, *Forensic Methods and Procedures Applied to Child Custody Evaluations: What Judges Need to Know in Determining a Competent Forensic Work Product*, 38(2) JUV. & FAM. CT. J. 21 (2000); Jonathan W. Gould & Debra H. Lehrmann, *Evaluating the Probative Value of Child Custody Evaluations*, 53 JUV. & FAM. CT. J. 17 (2002).

⁵⁰ Jonathan W. Gould, *Conducting Scientifically Crafted Child Custody Evaluations, Part One: A Model for Interdisciplinary Collaboration in the Development of Psycholegal Questions Guiding Court Ordered Child Custody Evaluations*, 37 FAM. & CONCILIATION CTS. REV. 64 (1999); Jonathan W.

ator clearly defines the questions and areas of concern that guide the entire evaluation. This approach has been termed a “minimalist approach,” because it provides focused support and information for the trier of fact in resolving a child custody dispute.⁵¹ Thus, an important responsibility of the child custody evaluator is to take the legally relevant dimensions that are the court’s concern and to define each dimension in a manner that allows for proper psychological assessment. This results in a reliable child custody evaluation. Moreover, the evaluator must also demonstrate that these psycholegal dimensions have an empirical foundation in the psychological literature, and that they are related to the questions that are before the court.⁵² This results in a relevant child custody evaluation.

The scope of a custody evaluation is properly determined by the court’s order. Court orders regarding CCEs vary widely in the extent to which they specify the questions to be addressed by the evaluator. Frequently, court orders simply direct the litigants to undergo an evaluation in accordance with a cited custody statute. Alternatively, many orders request a “psychological evaluation” of the parent-litigants and their children. Too often, child custody evaluators do not further query the court or the attorneys involved in the case regarding the specific concerns that led to the order for evaluation. This common lapse on the part of evaluators constitutes poor practice.⁵³ Moreover, it may reflect a lack of understanding that the proper role of a forensic specialist in assisting the trier of fact is to provide reliable psychological information that is relevant to the pending legal issue.

Judges and attorneys can greatly increase the utility of evaluations by crafting court orders that pose referral questions specific to each family. This practice increases the likelihood that evaluators will address matters of central importance to the litigation, and diminishes the likelihood that evaluators will address

Gould, *Conducting Scientifically Crafted Child Custody Evaluations, Part Two: A Paradigm for the Forensic Evaluation of Child Custody Determination*, 37 FAM. & CONCILIATION CTS. REV. 159 (1999).

⁵¹ Amundson et al, *supra* note 39, at 63-87.

⁵² Failing that, the evaluator must make clear the basis on which his or her opinions are given, so that the court may decide how much weight to give the opinion.

⁵³ Gould, *supra* note 50, at 183-195.

irrelevant issues that confuse the litigation and increase the cost of the evaluation. For example, in a case in which one party has made allegations of excessively harsh physical discipline against the other, useful referral questions might focus on issues such as the manner in which each parent disciplines the children, the manner in which the children respond to each parent's behavior management strategies, whether the children fear the parent against whom the allegations were made, and each parent's capacity to tolerate frustration.

In the context of defining the scope of the evaluation, problems of reliability refer to instances in which the evaluator uses unreliable methods or goes beyond the scope of his or her training and expertise in offering opinions to the court. For example, many child custody evaluators go beyond their expertise in offering opinions on issues such as the comparative educational quality of school districts, the quality of community life in a particular geographic area, or the benefits of certain financial arrangements. When evaluators engage in this practice, they are no longer properly testifying as experts.

Problems of relevance refer to instances in which the evaluator offers opinions about issues irrelevant to the pending legal issue, or fails to explain the relationship between the parties' observed capacities and the pending legal issue. A particularly problematic situation can arise when an evaluator offers opinions about issues that are both irrelevant to the pending legal issue and highly prejudicial. For example, consistent with their primary training as clinicians who diagnose and treat psychiatric disorders, many child custody evaluators routinely report psychiatric diagnoses for both parent-litigants. We maintain that it is poor professional practice for evaluators to report such diagnoses in the absence of any indication that the court has concerns about the parents' diagnostic status, and in the absence of any demonstration how these diagnoses impact the litigants' capacity to parent. Placing a child in the primary custodial care of a parent suffering from "Generalized Anxiety Disorder and Personality Disorder Not Otherwise Specified, with avoidant and obsessive-compulsive features" sounds almost negligent—despite the fact that this diagnosis may have nothing whatever to do with care giving capacity.

B. *Use Forensic Interview Techniques*

A CCE represents an assessment of two parents' comparative custodial strengths and limitations. Thus, the evaluator must gather interview information regarding the litigants' parenting competencies according to variables that are directly related to the pending legal issue. This includes gathering information regarding, among other relevant variables, the parents' respective capacities to nurture, support, discipline, instruct, and supervise their child. In this manner, the evaluator will be able to compare and assess the parents' responses to interview questions that are directly relevant to the court's concerns. Moreover, the evaluator must make direct comparisons of similar sets of parental competencies to reach conclusions that will be of use to the court.

The child custody evaluator ensures the relevance of the interviews by asking questions that directly relate to the pending legal issue. According to the functional approach to CCEs,⁵⁴ a parent's past and present caretaking abilities are of greater relevance than distal events that do not relate to parenting. For example, in a case in which there are concerns regarding one parent's alleged propensity for harsh corporeal discipline, the evaluator should interview both parents about their disciplinary styles, behavior management strategies, methods of dealing with frustration, and beliefs regarding concepts such as obedience and deference to authority. Issues such as a parent's work history or prior romantic relationships may indeed have bearing on a given case, and we are not suggesting that evaluators should refrain from addressing topics not directly related to childcare. However, such issues should be evaluated within a prevailing discussion of functional parenting competencies. The evaluator's primary interviewing goal should be to elicit information from both parents regarding their caregiving strengths and limitations.

The child custody evaluator ensures the reliability of the interviews by gathering a reasonably uniform set of interview data, and by examining a parent's responses for consistency across various sources of parenting information. For example, a parent's responses can be examined for consistency across time (i.e., a

⁵⁴ See, e.g., THOMAS GRISSO, *EVALUATING COMPETENCIES: FORENSIC ASSESSMENTS AND INSTRUMENTS* (1986).

comparison of the parent's responses to identical questions given at different times), across question formats (i.e., a comparison of the parent's responses to structured and unstructured questions), across parties (i.e., a comparison of the parent's responses to the responses given by the other parent, and/or by the children), and across collateral sources (i.e., a comparison of the parent's responses with the responses given by collateral sources).⁵⁵ Returning to the case in which allegations have been made regarding one parent's excessively harsh physical discipline, the evaluator can examine each parent's responses to questions of disciplinary and behavior management strategies for consistency across a number of different sources of information. For example, one parent's vociferous denial of allegations of excessive physical punishment may be contradicted by interview data generated by the other parent, by the children, by a variety of collateral sources, and by the allegedly abusive parent's own responses to interview questions regarding parenting values and beliefs, disciplinary strategies, and frustration tolerance.

When reviewing a CCE, it is also important to assess the weight assigned by the evaluator to the interview data. In distinguishing between therapeutic and forensic roles, Stuart Greenberg and Daniel Shuman state that forensic evaluators must use a higher level of scrutiny for interview information than do therapists.⁵⁶ They note that forensic interviews take place in a significantly different context than do therapeutic or diagnostic interviews. In a therapeutic interview, the patient perceives that there is benefit to providing accurate and detailed information about his or her emotional condition, with the expectation that such information will assist treatment. In this context, the patient is motivated to provide the therapist with information that is as truthful as possible, even if the patient harbors concerns that such information might create an unfavorable impression. In contrast, in a child custody interview, the parent perceives that there is benefit to providing complimentary information about

⁵⁵ ROGERS, *supra* note 45, at 113 & 367; Richard Rogers, *Structured Interviews and Dissimulation*, in *CLINICAL ASSESSMENT OF MALINGERING AND DECEPTION* (Richard Rogers ed., 1988).

⁵⁶ Stuart A. Greenberg & Daniel W. Shuman, *Irreconcilable Conflict Between Therapeutic and Forensic Roles*, 28 *PROF. PSYCHOL.: RES. & PRAC.* 50 (1997).

his or her parenting, with the expectation that such information will assist his or her legal goal. In this context, the parent is motivated to supply the evaluator with caregiving information that is as favorable as possible, and to withhold information that might create an unfavorable impression of his or her parenting. Although it is possible that a given parent is being truthful and not attempting to influence the interview, it is more likely that the parent is attempting to positively influence the evaluator's perception of the parent's caregiving capabilities. It is important to remember that mental health professionals are not more skilled than are laymen at assessing the credibility of interview statements.⁵⁷ Therefore, in a forensic context, evaluators should assess the utility of interview data by comparing them to other sources of information, and by searching for general trends and consistencies across multiple data sources.

One way that a child custody evaluator can increase the reliability and relevance of interviews is by using a questionnaire that asks parents a standard set of questions, while also providing for opportunities to ask questions regarding areas of functioning that may be unique to only one parent. This method is referred to as a semi-structured interview format.⁵⁸ Unlike a fully structured interview format, which consists of a fixed set of questions that allow no opportunity for digression, a semi-structured format is sufficiently flexible to permit exploration of topics that are not predetermined but that may be of substantial importance nonetheless. Unlike an unstructured interview format, which lacks any predetermined questions, a semi-structured format is sufficiently methodical to permit the evaluator to ask the same set of general questions to each parent, while also permitting deviation from those questions into areas unique to that particular parent and his or her relationship with the child.⁵⁹ In this way, while collecting a set of data common to both parents, the evaluator can pursue additional areas that are specific to one parent or to the context of the evaluation. Therefore, the use of a semi-structured interview protocol provides a systematic and sci-

⁵⁷ PAUL EKMAN, *TELLING LIES* (1992).

⁵⁸ See generally, GRISSO, *supra* note 54 (credited with developing the semi-structured interview techniques); SCHUTZ, ET AL., *supra* note 45.

⁵⁹ Gould, *supra* note 50, at 167; Gould & Lehrmann, *supra* note 49, at 17-29.

entific means of data collection that is consistent with evidentiary requirements for scientific information.⁶⁰

When evaluating the quality of interviews in the context of a CCE, attorneys should be attentive to problems of reliability and relevance. Problems of reliability refer to the evaluator's gathering of interview data in a subjective, partial, or unscientific manner. One example of a reliability problem is "confirmatory bias," or an evaluator's tendency to seek out data that supports his or her preconceived hypothesis, and to ignore data that is inconsistent with that hypothesis.⁶¹ Confirmatory bias can significantly distort the reliability and utility (validity) of interview data, and can lead the evaluator to inaccurate or one-sided conclusions unsupported by other evidence.

Gould provides an excellent example of how confirmatory bias can diminish the reliability of interview information in a CCE: The evaluator directed the judge to pay particular attention to the father's interview data. It was highly credible, she testified. The father had been living with his 17-year-old son for about a year. Each had a history of relationship difficulties with the mother. The evaluator interviewed the father and son over 3 days. She concluded that the father and son had an accurate view of the mother. Their opinions were judged to be credible and consistent. Based solely upon the information drawn from the father's and son's interviews, the evaluator concluded that the mother was abusive and therefore a threat to her children. Custody of all three boys was recommended to the father.

When the mother was interviewed, the evaluator began by complimenting her son and former husband for teaching her so much about their family life. She followed this statement with asking the mother, "How long have you been abusing your children?" In the body of the report, the evaluator commented that as the interview with the mother continued, the mother appeared to become increasingly defensive and unwilling to provide detailed answers. The evaluator never saw how her opening com-

⁶⁰ Gould & Bell, *supra* note 49, at 21-27; Jonathan W. Gould & Philip M. Stahl, *The Art and Science of Child Custody Evaluations: Integrating Clinical and Mental Health Models*, 38 FAM. & CONCILIATION CTS. REV. 392 (2000).

⁶¹ Randy Borum, Randy Otto, & Stephen Golding, *Improving Clinical Judgment and Decision Making in Forensic Evaluation*, 21 J. PSYCHIATRY & L. 35 (1993).

ments frightened the mother, creating a cognitive set of caution and distrust. Based on the mother's defensive responses, the evaluator concluded that the child's credibility was firmly established.

When an independent evaluator became involved and gained access to collateral information, it became clear that the younger children's teachers, coaches, youth minister, therapist and neighborhood parents described mother-child interactions as above average. These same sources described the father-child and father-mother interactions as significantly problematic. The father was a weekend alcoholic who often became violent and verbally abusive. Police records showed three arrests for DUI and one court appearance for disorderly conduct. The seventeen-year-old son had also been drinking for about eighteen months and developed a style of verbal abuse similar to his father's. He had his license suspended until he was twenty-one for driving while under the influence. Father and son often drank together during the evenings.

Furthermore, test data revealed a father whose scores were significantly elevated on a number of scales suggestive of severe psychopathology. Elevation on each scale was tied directly to collateral data supportive of the father's substance abuse, disregard for rules, and highly suspicious beliefs.

Finally, interview data from the younger children revealed children who were afraid of their father, particularly when he was drinking. In separate interviews, they reported that their older brother often hit them and verbally abused them while the father was in the kitchen observing their interactions, sipping a beer and doing nothing to intervene.⁶²

Problems of relevance refer to the evaluator's gathering of interview data in a manner that fails to address the pending psycholegal issue of comparative parenting capacity. One example of a relevance problem is the use of a traditional "clinical interview" in the context of a CCE. The primary purpose of a clinical or diagnostic interview is the identification of psychopathology and emotional distress. An additional purpose is the identification of intervention or treatment methods most likely to facilitate the subject's recovery. Unless the court will evaluate an

⁶² Gould, *supra* note 50, at 74-75.

issue regarding a parent's diagnostic status or psychological well-being, clinical data regarding psychopathology is not relevant to the pending legal issue. Child custody evaluators who engage in traditional clinical interviewing are not only likely to fail to adequately address the pending legal issue, but are also on a "fishing expedition" for psychopathology that can lead them astray from the court's need for reliable and relevant information.

C. Psychological Testing with Objective and Self-Report Measures

The purpose of psychological testing in CCEs is to provide the court with a set of objective scientific data. Psychological tests can provide a reliable and valid set of data that allow for more precise measurement of individual characteristics than can be obtained from interviews alone. Along with data generated by other sources of information, psychological test data can provide objective support to an expert's opinion and produce data grounded in empirical research.⁶³ Moreover, the objective data generated by psychological tests can balance the bias and potential errors inherent in clinical interview data.⁶⁴ By using multiple tests, the evaluator can search for trends across the data, and can cross-check his or her hypotheses. By incorporating multiple measures of multiple dimensions of functioning, the evaluator can gather a wide range of information with which to understand each parent's comparative caregiving strengths and limitations, both as compared to one another and as compared with a group of peers.⁶⁵

In his analysis of the role of psychological testing in forensic evaluation, Kirk Heilbrun lists seven criteria that should be met by assessment instruments used in a forensic context.⁶⁶ Heilbrun states that tests used in forensic mental health assessment should be: (1) commercially available, adequately documented in tech-

⁶³ DAVID L. SHAPIRO, *PSYCHOLOGICAL EVALUATION AND EXPERT TESTIMONY: A PRACTICAL GUIDE TO FORENSIC WORK* (1984).

⁶⁴ FORENSIC APPLICATIONS OF THE MMPI-2 (Yossef S. Ben-Porath, et al., eds., 1995); JOSEPH T. McCANN & FRANK J. DYER, *FORENSIC ASSESSMENT WITH THE MILLON INVENTORIES* (1996).

⁶⁵ Gregory J. Meyer, et al., *Psychological Testing and Psychological Assessment: A Review of Evidence and Issues*, 56 AM. PSYCHOLOGIST 128 (2001).

⁶⁶ Heilbrun, *supra* note 45, at 257-272.

nical manual, and reviewed in at least two professional sources; (2) sufficiently reliable; (3) relevant to the ultimate legal issue, or to a psychological construct underlying the ultimate legal issue; (4) administered in a standard manner; (5) applicable to the population being assessed; (6) evidencing an objective test format with an actuarial basis for diagnosis and prediction (as opposed to a subjective or impressionistic interpretive method); and (7) amenable to the explicit assessment of response style.

Essentially, these seven criteria demand that assessment instruments used in a forensic context be both reliable (i.e., valid) and relevant.⁶⁷ Important considerations in choosing a psychological test include published psychometric data supporting its reliability and validity, its acceptance as scientific evidence in other jurisdictions, its relevance to the psycholegal questions being examined, and its basis in scientific theory.⁶⁸ An additional consideration is that the assessment instrument should be a tool generally relied upon by professionals in the field for use in child custody matters.⁶⁹ Finally, the test should generate hypotheses that are directly relevant to the psycholegal questions posed by the court. For example, in the assessment of parental competencies, several personality tests are available that may be used to generate hypotheses about whether measured personality features influence an individual's parenting. However, it is important to note that no personality tests measure parenting competency, nor has any constellation of personality traits been

⁶⁷ Forensic assessment is premised upon the idea of convergent validity. A critical issue in the use of a multi-trait/multi-method model is the extent to which distinct assessment methods provide unique versus redundant information. It is important to recognize that more data does not always mean more accurate results. For example, when using tests that are intercorrelated, it is possible that the predictive power of the two tests combined is less than the predictive power of the psychometrically sounder test. The idea behind multi-trait/multi-method assessment is to increase the predictive power of the data. If a method does not increase the ability of the data to predict the behavior of interest, then there is no incremental validity derived from the use of the method. If there is no incremental validity, then there is no reason to administer the test.

⁶⁸ FORENSIC APPLICATIONS OF THE MMPI-2, *supra* note 64, at 117-125; McCANN & DYER, *supra* note 64; KENNETH S. POPE, K.S., JAMES N. BUTCHER, & JOYCE SEELEN, THE MMPI, MMPI-2 & MMPI-A IN COURT: A PRACTICAL GUIDE FOR EXPERT WITNESSES AND ATTORNEYS (2d ed., 2000).

⁶⁹ Ackerman & Ackerman, *supra* note 46.

linked to skill as caregiver.⁷⁰ It is impossible to determine from test results alone if a parent's measured response patterns are related, either directly or indirectly, to parenting competencies.⁷¹

When discussing test results, child custody evaluators must remember that whether the test results accurately describe a parent is dependent upon the degree to which other sources of information provide confirming or disconfirming data. Current forensic practice is to frame psychological test interpretations as hypotheses or general trends, and to avoid considering test results in isolation.⁷² Similarly, current forensic practice is to describe interpretative statements as actuarial and expert predictions based upon test results. Personality test results may indicate that a parent exhibits characteristics similar to individuals with similar response patterns; however, such test results are probabilistic in nature. Moreover, evaluators should interpret test results cautiously and in light of other data collected from multiple sources.⁷³ It is critical that the evaluator understand that test results provide only hypotheses, which then must be subjected to verification from alternative data sources.⁷⁴

⁷⁰ Grisso, *supra* note 54 ; McCANN & DYER, *supra* note 64.

⁷¹ SCHUTZ ET AL., *supra* note 45.

⁷² ELIZABETH M. ELLIS, *DIVORCE WARS: INTERVENTIONS WITH FAMILIES IN CONFLICT* (2000).

⁷³ Gould, *supra* note 50.

⁷⁴ In *In re B.M.*, 682 A.2d 477, 481 (Vt. 1996), a termination proceeding, the Vermont Supreme Court addressed the limitations of testing in a forensic assessment of parenting capacity:

[T]he court's emphasis on psychological testing is disturbing. Such tests, when relied on by expert witnesses, may have a small place in the overall evaluation of a person's parenting ability. Parents facing the loss of parental rights, however, must be judged on their conduct, not on their test-taking skills or psychological traits. In this case, for example, the court first labeled father, based on his MMPI results, as 'a person with hedonistic, narcissistic and impulsive tendencies and over controlled hostilities.' The Court then linked these personality traits to likely behaviors, noting that 'such, persons typically seek immediate gratification, blame others for their own problems, and manipulate others for their own desires, experiencing little guilt about the effects of their actions on other.' Finally, the court closed the door on possible changes or improvements, finding that 'these are consistent and pervasive traits which no form of intervention will change.' Although recognizing that parenting skills can be learned, the court nonetheless found that 'in times of stress [father] will fall back on his

Whether the tests are accurate enough for such use is an open question; and such tests raise concerns of cultural, educational, and socioeconomic bias. Moreover, characterizing individuals as bad parents based on “pervasive traits which no form of intervention will change” is inconsistent with the goal of fostering parental improvement. Finally, the use of psychological tests in a forensic evaluation should include a discussion of the limitations of the test data.⁷⁵

Examination of the scientific integrity of the measurement tools used in CCEs goes to the heart of the question of reliability. If an evaluator elects to use a given measurement tool in a CCE, the tool should evidence an appropriate level of scientific reliability and validity with regard to the specific issue in dispute. Such psychometric information allows the evaluator to consider the strength and limitations of the test data; without such information, evaluators have little ability to gauge the accuracy of the data upon which their conclusions are based. If a test used to measure a factor does not have adequate reliability, then the data upon which the evaluator’s interpretations, conclusions, and recommendations are based will be seriously flawed.⁷⁶ Moreover, if an evaluator elects to use a given measurement tool in a CCE, the evaluator should provide information about whether the instrument in question has normative data for male and female custody litigants (and, if so, how each parent’s scores com-

inherent personality traits.’ The court was apparently further persuaded by the correlation of father’s test results, finding that the ‘PASS results and MMPI scores reinforce each other and show a pervasive lack of empathy.’ We are unable to share the court’s confidence in this fact, as the findings and the record lack a meaningful explanation of the purposes, appropriate uses, or scoring methods for these tests. For example, expert testimony and the court’s findings emphasize that father’s PASS results were ‘clinically low.’ According to the PASS manual, however, the PASS is scored subjectively, by the individual evaluator. B. BRICKLIN, PARENT AWARENESS SKILLS SURVEY MANUAL 6 (1990). There is no evidence in the record to explain the expert’s scoring decisions or standard for comparison, or to justify reliance on the test scores in a proceeding to terminate parental rights. . . . The PASS manual provides little help in understanding the test scores, however, because the scoring is subjective. . . .

⁷⁵ Committee on Ethical Guidelines for Forensic Psychologists, *supra* note 7; American Psychological Association, *supra* note 40.

⁷⁶ Gould & Lehrmann, *supra* note 49.

pare to such normative data). An increasing pool of empirical data is developing with regard to how male and female custody litigants score on a number of psychological tests commonly used in CCEs.⁷⁷

In addition to demonstrating reliability, psychological tests ought to yield data that are relevant to the issues of concern to the court. In other words, psychological tests used in CCEs should demonstrate a valid scientific connection to the pending legal issue. For example, many evaluators use the Minnesota Multiphasic Personality Inventory - Second Edition (MMPI-2) and the Millon Clinical Multiaxial Inventory - Third Edition (MCMI-III) in child custody evaluations.⁷⁸ Although neither test directly measures parenting capacity, the inferences drawn from these tests may provide useful information about a parent's personality characteristics and emotional style. These constructs are relevant in the context of a CCE because a parent's psychological functioning is relevant to the issue of comparative custodial suitability. Normative data regarding the male and female custody litigants have now been published⁷⁹ that enable evaluators to compare a given parent's scores on the MMPI-2 and MCMI-III to normative scores obtained by other custody litigants, thereby further increasing the relevance of the obtained data.

In the context of psychological testing, problems of reliability refer to the use of measurement tools that lack the requisite scientific reliability and/or validity. One example of a reliability problem is the use of projective drawings to make inferences about a parent's psychological functioning, or about the issue of comparative parenting ability. Projective drawings lack the necessary validity and reliability for admissibility in court.⁸⁰ At a minimum, the subjectivity of the administration and interpretation procedures for projective drawings render them inappropriate for use in a legal proceeding. No normative data exist regarding the personality correlates of projective drawings among adults. Similarly, no empirical behavioral science litera-

⁷⁷ See, e.g., Bathurst et al, *supra* note 46; Joseph T. McCann, et al., *The MCMI-III in Child Custody Evaluations: A Normative Study*, 1 J. FORENSIC PSYCHOL. PRAC. 27 (2001).

⁷⁸ Ackerman & Ackerman, *supra* note 3, Bow & Quinnell, *supra* note 1.

⁷⁹ Bathurst et al, *supra* note 46; McCann, et al., *supra* note 77.

⁸⁰ Faigman (Evidentiary Status), *supra* note 45.

ture exists demonstrating that projective drawings are related to any specific element of a parent-child relationship, or are predictive of any particular parenting practices or developmental outcomes. It therefore constitutes poor professional practice for an evaluator to render psycholegal conclusions about adult personality structure and psychological functioning on the basis of projective drawings.

The use of psychological tests that do not provide data that is related, either directly or indirectly, to the pending legal issue raise issues of relevance. For example, many child custody evaluators administer measures of intellectual functioning to each parent.⁸¹ However, unless specific concerns arise regarding a parent's intellectual functioning, such data will be irrelevant to the evaluation. In other words, a psychological test ought to provide data that is useful in answering some question or issue before the court. Therefore, unless parental intellectual functioning is an issue before the court, there is no psycholegal reason to administer such a test. Significantly, there exists no empirical data to suggest that parents of above-average intellectual functioning provide more competent parenting than do parents of average intellectual functioning. Similarly, there exists no empirical data that links above-average intellectual functioning with superior caregiving skills or with any uniquely positive elements of a parent-child relationship. One parent's superior scores on a test of intellectual functioning might inadvertently give the court the incorrect impression that the "smarter" parent will make the "better" parent. In such an instance, the presentation of data that appears to suggest a scientific comparison but does not will be more prejudicial than probative.

Once an assessment tool is identified in a report as a psychological test, and if the case is in a state in which *Daubert* is accepted, the expert should be prepared to defend the choice and the use of a particular test in a *Daubert* challenge. Such a challenge may include explaining to the judge information about a test's underlying theory of science, whether it has been published in a peer-reviewed journal, status as a current standard among

⁸¹ Ackerman & Ackerman, *supra* note 46, at 579-84.

the psychological community, psychometric data about its reliability, validity, base rate estimates and its falsifiability.⁸²

If the expert bases his or her opinion on a certain instrument, that instrument's scientific validity for the purpose used may be the focus of a *Daubert* challenge. If the expert bases his or her opinion on a particular methodology, then it is the methodology that will be the focus of a *Daubert* challenge.

With respect to psychological testing, there are two ways to approach a *Daubert* challenge: First, focus attention on each test used in the battery of tests. Most of the psychological tests and measures used in child custody evaluations might have difficulty surviving a *Daubert* challenge because the tests have not been developed for use in child custody assessments. Therefore, the expert should not base his or her opinion upon the results of a specific test or a specific set of tests.

Second, the expert should focus less attention on individual tests and more attention on the scientific methodology used in the evaluation process. Rather than examining each test used in the evaluation process, the evaluator describes the usefulness and breadth of data from alternative independent sources of information. As noted in *Daubert*: "Scientific methodology today is based on generating hypotheses and testing them to see if they can be falsified; indeed, this methodology is what distinguishes science from other fields of human inquiry."⁸³

The evaluator must describe how information from any one data source may be used to generate hypotheses about the parent, the child, or the family. It is important to describe how information from one source of data is used to confirm or disconfirm hypotheses generated from other independent sources of information. One view is that psychological test data are weighed no more heavily than collateral data or direct observational data. Another view is that psychological testing should be weighed more heavily than other sources of data because a well-developed test will have standardized norms, quantitative measurements, multiple validity studies, standardized observational conditions and other psychometric properties.

⁸² David Medoff, *The Scientific Basis of Psychological Testing: Considerations Following Daubert, Kumho and Joiner*, 41 FAM. CT. REV. 199 (Apr. 2003).

⁸³ *Daubert*, 509 U.S. at 593.

How the usefulness of data is weighed from different information sources depends on the characteristics of each case. It is the convergence of independent sources of data that help to make one hypothesis more likely than another to be supported, and the weight assigned to each information source may vary from case to case depending upon the quality of the data and the way in which the data may be interpreted within the larger context of the family system.

In reality, *Daubert* challenges are rarely used in child custody cases. If it is a *Frye* jurisdiction, the task of the lawyer may be very different in defending the use of a psychological test or measure. Several publications describe how psychologists and other mental health professionals use psychological tests in custody evaluations,⁸⁴ and a number of upcoming articles investigate evaluators' knowledge of *Frye* and *Daubert* issues when selecting and interpreting psychological tests.⁸⁵ Based upon a *Frye* standard, it is possible that the use of the House-Tree-Person test, Sentence Completion test or Thematic Apperception Test would be ruled admissible because of their general acceptance among evaluators as reported in these peer-reviewed articles.

It is the intention of the current APA Ethics Code and the Specialty Guidelines for Forensic Psychologists to focus attention on the reliability, validity and relevance of tests and measures used in a forensic context. A test that is widely used by colleagues does not mean that the test is psychometrically sound.

An assumption built into the *Frye* standard is that a test would not be commonly used among professionals in a field if its reliability had not been previously demonstrated. As summarized previously in a recent Illinois State Supreme Court decision, "A technique, however, is not 'generally accepted' if it is experimental or of dubious validity. Thus, the *Frye* rule is meant to exclude methods new to science that undeservedly create a perception of

⁸⁴ See, e.g., Ackerman & Ackerman, *supra* note 46; Bow & Quinnell, *supra* note 1. .

⁸⁵ James N. Bow, James R. Flens & Jonathan W. Gould, J.W., *Testing in Child Custody Evaluations- Selection, Usage, and Daubert Admissibility*, ___ J. FORENSIC PSYCHOL. PRAC. (forthcoming 2005); James N. Bow, James R. Flens & Jonathan W. Gould, *An Analysis of Administration, Scoring, and Interpretation of the MMPI-2 and MCMI-III in Child Custody Evaluations* (on file with authors).

certainty when the basis for the evidence or opinion is actually invalid.”⁸⁶

Such is not the case with many assessment techniques used in the child custody field. Many clinical assessment techniques drawn from clinical practice have been used in child custody evaluations and, until recently, have been accepted as commonly used among custody evaluators despite their lack of reliability and relevance.⁸⁷

The use of unreliable, clinically derived techniques such as human figure drawings, sentence completion tests and other similar projective techniques may have earned a place in the scientific community and have undeservingly created a perception of reliability when, in fact, there is no basis for such belief. An assumption of the *Frye* standard that a principle or technique is not generally accepted in the scientific community if it is by nature unreliable does not appear to be reflected in custody evaluators’ historical use of projective techniques. Unreliable methods and the interpretation of unreliable information that has been drawn from those unreliable methods have been used as a basis upon which evaluators have offered opinions about custodial placement and visitation access. Reliance upon unreliable techniques serves neither the families nor the courts. Fortunately, child custody evaluators increasingly focus on the use of reliable and relevant assessment techniques.⁸⁸

D. *Direct Behavioral Observations of Parent-Child Interactions*

When a forensic evaluator is assessing a parent’s caregiving capacities, the evaluator must engage in direct observation of parent-child interactions.⁸⁹ This is true regardless of whether the evaluator is assessing child custody, parental competency, or parental risk to the child and is expressly stated in the professional

⁸⁶ *Donaldson v. Central Illinois Pub. Serv. Co.*, 767 N.E.2d 314, 324 (Ill. 2002).

⁸⁷ *Ackerman & Ackerman*, *supra* note 46, at 565-67.

⁸⁸ *Bow & Quinnell*, *supra* note 1, at 261-68.

⁸⁹ Some commentators suggest that such observations are not always necessary for older children involved in custody disputes. *See, e.g.*, *STAHL*, *supra* note 43. We are not in agreement with this position.

guidelines.⁹⁰ For example, the Guidelines for Child Custody Evaluations in Divorce Proceedings state that a child custody evaluation should include “an evaluation of the interaction between each adult and child.”⁹¹ Direct behavioral observation can provide information about a parent’s caregiving strengths and weaknesses, such as a parent’s communication skills, perception of the child, provision of structure and support, manner of expressing love, manner of providing discipline and knowledge and expectations regarding the child’s developmental needs and abilities. Since the heart of any CCE is the relationship between parent and child, direct behavioral observation increases the relevance of the data. Moreover, since observation provides an opportunity to test hypotheses regarding parental strengths and weaknesses, direct behavioral observation increases the reliability of the data.

Observation of parent-child interactions can occur in structured or unstructured formats. Structured observational formats typically require a parent and child to engage in a series of tasks, or require an observer to score the parent and child according to a series of interactional ratings. For example, the Parent-Child Early Relational Assessment⁹² is a structured parent-child observation system that can be used to assess the quality of the relationship between a young child and his or her parent. The parent and child are observed interacting during four segments of an observational protocol: a feeding exercise; a structured task (e.g., reading); free play; and a separation and reunion.⁹³ Un-

⁹⁰ American Psychological Association, *supra* note 40; Association of Family and Conciliation Courts, *supra* note 45; Committee on Professional Practice and Standards, APA Board of Professional Affairs, *Guidelines for Psychological Evaluations in Child Protection Matters*, 54 AM. PSYCHOLOGIST 586 (1999); SCHUTZ ET AL., *supra* note 45.

⁹¹ American Psychological Association, *supra* note 40, at 678.

⁹² ROSEANNE CLARK, THE PARENT-CHILD EARLY RELATIONAL ASSESSMENT (1985); Roseanne Clark, Andrew Paulson & Susan Conlin, *Assessment of Developmental Status and Parental-Infant Relationships*, in HANDBOOK OF INFANT MENTAL HEALTH (Charles Zeanah ed., 1993).

⁹³ An evaluator may indicate that he or she has followed a standardized procedure to gather observational data. Although some of these standardized procedures have an empirical basis, some have little or not empirical evidence to support their validity. Moreover, some empirically-based observational procedures require that a practitioner undergo extensive training to qualify as an expert in their use (e.g. the “strange situation,” which is used to classify young

structured parent-child observation consists of simply observing the parent and child together in an office setting or a naturalistic environment such as the home, where observations can be made of daily activities such as meals and bedtime routine.⁹⁴ In practice, Marc Ackerman and Andrew Kane are likely correct in their assertion that “[t]here are probably as many different ways to perform observations as there are evaluators.”⁹⁵

Regardless of the setting and structure for the observation, the evaluator’s task is to assess the nature and quality of the interactions between the parent and child, including assessing such critical issues as the way in which the parent expresses love, affection or support to the child; the way in which the parent provides discipline or structure to the child; the parent’s capacity to understand and respond to the child’s cues; the parent’s capacity to experience the child as a separate individual, with distinct preferences and relationships; the parent’s sense of enjoyment or competence in the parenting role; and the parent’s sense of frustration or inadequacy in the parenting role.

Careful observation of parent-child interactions can be significantly revealing, as the evaluator has the opportunity to witness patterns of interaction outside the parent’s awareness. For example, phenomena such as a parent’s sour expression when the other parent is mentioned or a parent’s repeated depreciation of the child in an attempt to be educative may become obvious during an observation session. Even more important, the child’s responses to parental behavior also become clear. An example follows: During interviews, Ms. Jones seemed overly concerned with her four-year-old son’s cognitive development to the exclusion of almost all other aspects of the child’s development. Instructed by the evaluator to bring materials to the observation session with which she and her son would like to work, Ms. Jones brought in a large shopping bag full of reading material and games designed to teach arithmetic. She proceeded to invite the

children’s attachment status). Whenever an evaluator claims to be using a standardized observational method, the attorney should attempt to clarify the support the method has from validation studies, as well as the evaluator’s qualification to use the method.

⁹⁴ STAHL, *supra* note 43.

⁹⁵ MARC J. ACKERMAN & ANDREW W. KANE, PSYCHOLOGICAL EXPERTS IN DIVORCE ACTIONS 159 (1998).

child to work with them. To the evaluator's surprise, the child instantly cuddled up to his mother, who put her arm around him with warm affection. He worked assiduously at the games with evident delight. When he grew tired of an activity, he readily communicated this to his mother; his mother repeatedly responded by encouraging him to select an activity he would enjoy more or by gently telling him that the activity was difficult at this point in time but that she was sure he could do it if he persisted. Observing the mother and child together, the evaluator concluded that the mother's style, which might have been problematic for some children, worked extremely well for her son.

Problems of reliability can occur when an evaluator offers conclusions in the absence of any observational data. For example, many evaluators describe observational sessions as revealing "a warm relationship" or "a positive attachment," without elucidating the data upon which these conclusions are based. In the absence of any observational data to support this conclusion (e.g., warm physical contact, gentle redirection, verbal praise and encouragement), it is not possible to assess the reliability of the evaluator's conclusions. Problems of relevance can occur when an evaluator offers peripheral data that lack any relationship to the underlying psycholegal issue of comparative parenting capacity. For example, some evaluators report observational sessions as verbatim transcripts and do not appear to employ any guiding observational methodology. In the absence of a conceptual framework with which to organize and understand observational data, it is likely that the evaluator's conclusions will lack the requisite relevance.

Finally, videotaping behavioral observation sessions can facilitate both the reliability and relevance of the data gathered. Currently, the videotaping of such sessions is uncommon in most jurisdictions. Videotaping has some disadvantages, including cost, time and potential evidentiary problems. However, videotapes can provide a particularly informative record of parent-child interactions. Reviewing the videotape can allow the evaluator to observe exchanges that went unnoticed during the observation session itself. Moreover, videotaping permits third parties (such as the court) to see the data described in the evaluator's report, including gestures and vocal tones, which can be extremely significant. We believe that because videotaped record-

ings provide one of the best means of reducing reliance on impressionistic descriptions of parent-child interactions and thus increase the reliability of behavioral observations, legal professionals should encourage videotaping of observation sessions whenever practical.

E. *Interviews with Collateral Sources and Review of Relevant Records*

The acquisition of reliable and relevant collateral information is arguably the most important component of a child custody evaluation. Forensic evaluation differs from clinical evaluation in its emphasis on establishing historical truth.⁹⁶ Forensic evaluators can utilize collateral data sources to help determine the facts underlying the psycholegal issue before the court. Collateral data serves a number of important functions in a CCE. First, such data can support or contradict a custody litigant's allegations regarding comparative parenting competency. Second, such data can control for the potential effects of deception and malingering, since parents may—intentionally or unintentionally—distort information in a manner that serves their legal position. Third, such data can provide a way for the evaluator to increase the confidence of interpretations and conclusions, because the obtained information derives from sources external to the evaluation and provides external validation in support of one or more hypotheses. Fourth, such data can make a significant contribution to an understanding of the litigant's behavior prior to the current legal dispute. For example, in a child custody case, collateral interviews can provide historical data about the parent-child relationship that may be critical to the examiner, yet otherwise unavailable.

The decision to interview collateral sources should be “based upon criteria of relevancy, reliability and necessity.”⁹⁷ A more valuable collateral source will be one who is not related to either parent and who has no vested interest in the outcome of the evaluation. We support William Austin's model⁹⁸ of dia-

⁹⁶ Shuman, *supra* note 17.

⁹⁷ Herbert N. Weissman, *Child Custody Evaluations: Fair and Unfair Professional Practices*, 9 BEHAV. SCI. & L. 469, 473 (1993).

⁹⁸ William G. Austin, *Guidelines for Utilizing Collateral Sources of Information in Child Custody Evaluations*, 40 FAM. CT. REV. 177 (2002).

gramming collateral sources as a series of concentric circles in which more distant emotional relationships with the litigants produce more neutrality. According to this diagrammatic model, inner circles are occupied by family members and friends; circles of middle distance are occupied by individuals such as coaches and scout leaders; and outer circles are occupied by individuals such as pediatricians, teachers and other community professionals.

The reliability of collateral interviews is increased when the evaluator uses neutral, non-aligned sources who can provide a credible view of the litigants' parenting over time. The relevance of collateral interviews is increased when the evaluator gathers information about a parent's real-life caregiving practices, competencies, and difficulties.⁹⁹ Any competent CCE must include information about how the parent and child operate in the real world, outside the artificial and contrived circumstances of the evaluator's office. Obtaining information from people who have direct observational knowledge of the parent and child in different situations is often the most important data obtained in a CCE. Evaluators can also interview sources who do not have knowledge of parent-child interactions, but who have observational knowledge of the child's functioning and adjustment in a variety of domains (e.g., academic, social, emotional), such as teachers or day care providers.

Problems of reliability can occur when the evaluator interviews individuals who are aligned with one litigant and are therefore personally invested in the outcome of the litigation. Examples of such individuals include a litigant's parents, siblings or new spouse.¹⁰⁰ Although it will be helpful to talk with such

⁹⁹ *Id.*; see also Gould, *supra* note 50, at 162-63.

¹⁰⁰ It is also important for the evaluator to understand how each collateral informant may know each parent. For example, it is not uncommon for a teacher to have more contact with a stay-at-home mother than with a working father. When such situations arise, it is important for the evaluator to examine the degree to which a teacher, who otherwise would be considered a neutral informant, may have aligned herself with one parent due to her increased contact with that parent and that parent's perspective on the custodial conflict. Similarly, when interviewing a child's therapist, it is important for the evaluator to determine whether the therapist has received information from both parents before determining that the therapist is a neutral source of information. Therapists, who may be viewed as credible because of their professional credentials, may become aligned with their adult patient or with the parent of their child

individuals, it is critical that the evaluator treat their reports similarly to information from the parent-litigants themselves: as data that may be intentionally or unintentionally biased, and that therefore must be corroborated with other sources of information. The more emphasis an evaluator places on collateral informants from within the parent's close emotional circle, the greater the likelihood that the obtained information is biased. Conversely, when information from a parent's close friends and family is corroborated by information from neutral and non-aligned informants, an evaluator can have greater confidence in the accuracy and utility of the information.

Problems of relevance can occur when the evaluator accepts and considers information not relevant to the issue of comparative parenting capacity. For example, parents may ask their family and friends to provide letters that serve as "character affidavits" attesting to the parent's good will and moral character. When such letters are provided by friends or co-workers who have no observational knowledge of parent-child interactions and little or no knowledge of the child involved in the dispute, they lack the requisite relevance to be of use. There may be cases in which such information is of use, and we are not suggesting that evaluators should never consider input from friends or co-workers. However, in our experience, these letters rarely contain any information relevant to the issue of parenting capacity.

IV. Complex Issues in Child Custody Evaluations: Domestic Violence and Child Alienation

It is particularly important for legal professionals to assess the reliability and relevance of CCEs in complex cases that involve allegations of domestic violence, abuse or child alienation. The forensic assessment of such allegations remains a controversial topic in CCEs. These assessments present a unique challenge because of the complexity of psychological variables involved in a comprehensive assessment and because of the social policy im-

patient who supports treatment. Therefore, they may be an unreliable source of collateral information.

plications and political passions that are evoked when such allegations are part of a custody case.

Writing about allegations of domestic violence in the context of child custody litigation, Austin notes, "there is probably no forensic question on which overreaching by mental health professionals has been so common and egregious."¹⁰¹ Overreaching occurs because evaluators reach conclusions based upon inadequate or incomplete data, or upon outdated research or personal beliefs presented as professional judgments. Significantly, a relationship exists between allegations of child alienation and domestic violence. Recently, the concept of alienation has undergone both revision¹⁰² and critique.¹⁰³ Child alienation is currently viewed as a family system process in which the alienating parent, the alienated parent and the child all contribute to dysfunctional relationships within the family system.

One important criticism of CCEs that assess for domestic violence or child alienation is that evaluators often are poorly trained to discriminate alienation from abuse. A parent who has been abused, or who is protecting a child from abuse, may appear to be alienating the child from the abusive parent when, in fact, the protective parent is attempting to keep the child safe. As a result of this criticism, some researchers have re-conceptualized alienation by looking at the child's behaviors and attachments and assessing whether the parent's behaviors may be alienating in nature.¹⁰⁴ These researchers have proposed a series of areas to explore in determining whether abuse and/or aliena-

¹⁰¹ William G. Austin, *Assessing Credibility in Allegations of Marital Violence in the High Conflict Child Custody Case*, 38 FAM. & CONCILIATION CTS. REV. 462, 463 (2000).

¹⁰² Joan B. Kelly & Janet R. Johnston, *The Alienated Child: A Reformulation of Parental Alienation Syndrome*, 39 FAM. CT. REV. 249 (2001); Richard A. Warshak, *Current Controversies Regarding Parental Alienation Syndrome*, 19 AM. J. FORENSIC PSYCHOL. 29 (2001).

¹⁰³ Carol S. Bruch, *Parental Alienation Syndrome and Parental Alienation: Getting it Wrong in Child Custody Cases*, 35 FAM. L.Q. 527 (2001); Kathleen C. Faller, *The Parental Alienation Syndrome: What It Is and What Data Support It*, 3 CHILD MALTREATMENT 100 (1998).

¹⁰⁴ Leslie Drozd, Toby Kleinman, & Lenore Walker, *How To's in Family Law Cases Involving Domestic Violence*, in CONFLICT RESOLUTION, CHILDREN AND THE COURTS: 38TH ANNUAL CONFERENCE OF THE ASSOCIATION OF FAMILY AND CONCILIATION COURTS 309 (2001).

tion are present in a custody case. Among the questions to investigate are: (1) is there a problem with the child's attachments?; (2) is there a problem with the child's behavior?; (3) if so, is there a reality-based reason for the child's troubled behavior?; (4) are there reasons to believe that the child has been exposed to some form of abuse; (5) are there reasons to believe that the child has been the victim of some form of abuse?; (6) are there reasons to believe that the child has interpreted events as abusive?; and (7) if the child has been exposed to or a victim of abuse, is the abuse "pure abuse" or is it combined with alienation dynamics?

The competent evaluator needs to be aware of how different factors are empirically linked to specific areas of family functioning. The competent evaluator also needs to systematically explore each of the variables known to be associated with different forms of violence and maltreatment. For example, if a referral question focuses attention on partner violence during the marriage, the evaluator must explore concerns about child abuse perpetrated by each parent, and not merely by the alleged aggressor.

V. Considerations in Using Mental Health Professionals in Child Custody Litigation

Given the above analysis of methodological problems common to CCEs, a threshold question in the preparation of a custody dispute case is whether to use a forensic mental health expert at all. In cases in which the facts clearly mitigate in favor of one parent and against the other as custodian, expert testimony may add little or no new information. Conversely, a CCE will be indicated in cases in which expert testimony can assist the trier of fact in understanding the evidence or in determining the facts at issue and when the expert is qualified by adequate knowledge, skill, experience, training or education.¹⁰⁵

If lawyers are generally ill-prepared to cope with scientific and technical material, many mental health professionals are sim-

¹⁰⁵ While a licensed mental health professional may generally be permitted to opine in all areas of his or her discipline, when a matter arises in a new and emerging field that is highly specialized, most courts also require additional expertise before admitting testimony on the subject. Examples of such highly specialized areas include clinical phenomena such as recovered memories of sexual abuse and clinical diagnoses such as Munchausen's Syndrome by Proxy.

ilarly unprepared to work with legal concepts and evidentiary issues. Nonetheless, forensic training is more available in the mental health field than is behavioral science and research training in the legal field. The preferred custody expert is one who has had significant education in the area in which he or she is opining, as well as experience with the rules of evidence.¹⁰⁶ Review and verification of the expert's curriculum vitae is essential; copies of prior forensic mental health reports and authored articles may also prove useful. If psychological tests will be used, the expert can be asked for a copy of the manual and any articles in the legal and scientific literature that explain or critique the tests and their application in CCEs.

Local rules will govern discovery opportunities with respect to adverse expert testimony. To the extent permitted by local law, production can be demanded of all notes and material in any media (including electronic media), and any test results (including raw data). Practitioners should demand an index of any material asserted to be privileged, the nature of the material, and the basis for the assertion of privilege. This can constitute the foundation for a subsequent motion to compel production or an in camera inspection of the material asserted to be privileged. Finally, since most law school curricula do not include statistics or research methodology classes—fields of learning essential to

¹⁰⁶ Who qualifies as an expert obviously depends upon the issues in a given case. Where a specific clinical condition is at issue, it would appear that training and experience relevant to the diagnosis and treatment of that condition would be necessary. However, this is not necessarily the case. Jurisdictions do not agree as to the necessity of a local license or specific training and experience. For example, North Dakota takes the position that Federal Rule of Evidence 702 does not require licensure in a particular field, or licensure in the court's jurisdiction, to qualify as an expert. Rather, it is the witness's actual qualifications that count. *Anderson v. A.P.I. Co.*, 559 N.W.2d 204, 206-07 (N.D. 1997); *State v. Carlson*, 559 N.W.2d 802, 809 (N.D. 1997); *Oberlander v. Oberlander*, 460 N.W.2d 400, 402, (N.D. 1990). Furthermore, in North Dakota, any educated and experienced psychologist should be able to qualify as an expert to testify about child custody factors. Unfamiliarity with the statutory factors affecting the legal determination of custody, a potential conflict of interest, and bias might affect the weight given the opinion, but these factors go to the credibility—not to the admissibility—of the evidence. *Kluck v. Kluck*, 561 N.W.2d 263, 266 (N.D. 1997). Obviously, a review of the local jurisdiction's views on these issues is necessary both in choosing an expert and in questioning an adverse expert.

the litigation of admissibility issues and the examination and cross-examination of scientific experts—counsel may wish to hire a consulting expert who can explain the scientific methodology underlying the expert’s opinions, critique flawed or unscientific methodologies and assist in preparing direct and cross examination.¹⁰⁷

VI. Ethical Principles of Psychologists and Code of Conduct Relevant to *Daubert*

The American Psychological Association’s Ethical Principles of Psychologists and Code of Conduct (Ethics Code) set forth enforceable rules for conduct as psychologists.¹⁰⁸ The Ethical Standards (ES) are not exhaustive: if conduct is not specifically addressed, it does not mean the conduct is ethical or unethical.

The newest version of the Ethics Code came into effect on June 1, 2003, and amended the 1992 version of that Code. The amended Code is designed to give the psychologist a greater ability to exercise professional judgment regarding the appropriate response to a variety of situations, in part by increased use of the terms like “reasonably,” “appropriate,” and “potentially,” and decreased use of “must” or “should.” The general directive is to do what a “reasonable psychologist” would do.¹⁰⁹

The Ethics Code applies only to psychologists’ activities that are part of their scientific, educational, or professional roles as psychologists (i.e., counseling, clinical, research, teaching, forensic activities), not their private conduct. Further, a finding that a psychologist has violated a provision of the Ethics Code is not intended to be a basis of civil liability against that psychologist. A psychologist who violates the Ethical Standards faces sanctions

¹⁰⁷ See also Barbara Ellen Handschu, *Tips, Strategies for Questioning Your Expert on Direct Exam*, 15 *MATRIM. STRATEGIST* 1 (Dec. 1997); James J. Jimmerson, *Cross-Examining an Opposing Expert at Trial: Preparation (Part I of II)*, 18 *FAIR\$HARE* 2 (May 1998); Stuart B. Walzer & Jan C. Gabrielson, *Strategic Cross-Examination in Commandments of Cross-Examination*, 3 *LITIGATION* 18 (Winter 1977).

¹⁰⁸ American Psychological Association, *supra* note 40 .

¹⁰⁹ MARC J. ACKERMAN, *CLINICIANS GUIDE TO CHILD CUSTODY EVALUATIONS* (1995); PHILIP M. STAHL, *CONDUCTING CHILD CUSTODY EVALUATIONS: A COMPREHENSIVE GUIDE* (Supp. 2005).

ranging from educational advisories to an expulsion from the APA (if he/she is a member).

Several Ethical Standards have specific relevance to the requirements set down by the Court in *Daubert*.

1. ES 9.01(b) – “Except as noted in 9.01(c), psychologists provide opinions of the psychological characteristics of an individual only after they have conducted an examination of the individual adequate to support their statements or conclusions. When, despite reasonable efforts, such an examination is not practical, psychologists document the efforts they made and the result of those efforts, clarify the probable impact of their limited information on the reliability and validity of their opinions, and limit the nature and extent of their conclusions or recommendations.”¹¹⁰

Exception – ES 9.01(c) – “When psychologists conduct a record review or provide consultation and an individual examination is not necessary for the opinion, psychologists can explain this and the sources of information on which they based their conclusions and recommendations.”¹¹¹

This Ethical Standard specifically addresses the importance of in-person evaluations of individuals about whom psychologists will offer a professional opinion. Under this standard, with few exceptions, psychologists must conduct individual examinations sufficient to obtain personal verification of information on which to base their professional opinions and refrain from providing opinions about the psychological characteristics of an individual if they themselves have not conducted an examination of the individual adequate to support their statement or conclusions.

2. ES 9.02(a) – “psychologists administer, adapt, score, interpret, or use assessment techniques, interviews, tests, or instruments in a manner and for purposes that are appropriate in light of the research on or evidence of the usefulness and proper application of the techniques.”¹¹²
3. ES 9.02(b) – “psychologists use assessment instruments whose validity and reliability have been established for use with members of the population tested. When such validity or reliability has not

¹¹⁰ American Psychological Association, *Ethical Principles of Psychologists and Code of Conduct, Ethical Standards* § 9.01(b) (2002) available at <http://www.apa.org/ethics/code2002.pdf>.

¹¹¹ *Id.* at § 9.01(c).

¹¹² *Id.* at § 9.01(a).

been established, psychologists describe the strengths and limitations of test results and interpretations.”¹¹³

A psychologist should ensure that a test has been validated for use with individuals of the age, gender, ethnicity, etc. of the client. Some of this information will come from the test manual. The psychologist also needs to be sufficiently familiar with the research on the test to be able to assess whether new research supports or questions particular uses of the test or interpretations of the results. If a test is used despite the lack of research-based support for the particular use, the psychologist is required to specify why it was used, the advantages of using it and any limitations on interpretations and recommendations as a result of its use.

4. ES 9.06 – “When interpreting assessment results including automated interpretations, psychologists take into account the purpose of the assessment as well as the various test factors, test-taking abilities, and other characteristics of the person being assessed, such as situational, personal, linguistic, and cultural differences, that might affect psychologists’ judgments or reduce the accuracy of their interpretation. They indicate any significant limitations of their interpretations.”¹¹⁴
5. ES 9.08(a) – “Psychologists do not base their assessment or intervention decisions or recommendations on data or test results that are outdated for the current purpose.”¹¹⁵

VII. Conclusion

Given the psycholegal import of CCEs, it is imperative that forensic mental health professionals tender evaluations that have

¹¹³ *Id.* at § 9.02(b). Of note, on April 8, 2005, a Youngstown, Ohio psychologist’s license was suspended for two years. Among the factors considered by the Board in making its decision was the use by the psychologist of “an insufficiently validated instrument . . . as a basis for reaching conclusions in psychological assessments.” Other factors were also considered by the Board, so there is no way of knowing how much weight was assigned to this one particular element. Also of interest, from the same case, the Board cited a section of Ohio law referred to as the “Specialty Standard of Care,” which states that “one who undertakes practice in a given specialty area will be held to the standard of care within that specialty while he/she is practicing in that area.” The suspension is stayed pending its appeal. See <https://license.ohio.gov>

¹¹⁴ *Id.* at § 9.06.

¹¹⁵ *Id.* at § 9.08(a).

been conducted with due regard for scientific methodology, behavioral science literature and ethical guidelines. Evaluators who purport to assess the best psychological interests of children involved in custody disputes must take precautions not to inadvertently harm those interests. As Chief Justice Frank D. Celebrezze of the Ohio Supreme Court wrote, “While statutes can be amended and case law can be distinguished or overruled, we take judicial notice of the fact that children grow up only once. When a mistake is made in a custody dispute, the harmful effects are irrevocable.”¹¹⁶

¹¹⁶ In re Wonderly, 423 N.E.2d 420, 427 (Ohio 1981).

