

Psychological Testing Can Be of Significant Value in Child Custody Evaluations: Don't Buy the "Anti-Testing, Anti-Individual, Pro-Family Systems" Woozle

by

Sol R. Rappaport,* Jonathon Gould,** and Milfred D. Dale***

Family courts making *best interests of the child* determinations frequently request that a mental health expert conduct a child custody evaluation. *Best interests* decisions are fact-intensive and involve multi-dimensional and multi-faceted assessments as part of creating an individualized parenting plan for each child and his or her family. Used within a multi-method approach to data gathering, psychological testing often helps evaluators develop hypotheses about the parties' behavioral tendencies, mental health issues, and psychological functioning as they may affect parenting, parent-to-parent communication, and other custody-related areas of concern.

Understanding how to properly utilize and apply information gained from psychological test data in child custody evaluations requires training, experience, and supervision in the forensic use of psychological tests. No single psychological test, or even battery of psychological tests, is capable of reliably replacing a court's consideration of numerous factors identified by state statutes and case law. This claim is unsupported by the relevant literature. Similarly dubious (and also lacking support in the literature) is the claim that competent child custody evaluators have been placing undue weight on psychological test re-

* Ph.D., ABPP. Dr Rappaport practices in Libertyville, Illinois.

** Ph.D., ABPP. Dr. Gould practices in Charlotte, North Carolina.

*** Ph.D., JD. Dr. Dale practices in Topeka, Kansas. Additional contributors to the article include William G. Austin, Ph.D., Leslie M. Drozd, Ph.D., James R. Flens, Psy.D., ABPP, Daniel Pickar, Ph.D., ABPP, and Allan Posthuma, Ph.D., ABPP.

sults without consideration of the multiple, independent sources of information that are gathered during a custody evaluation.

Several peer-reviewed articles¹ and texts² have explained in great detail the advantages and cautions to be considered in the use of psychological testing in child custody evaluations. Any proposal for a moratorium on use of psychological testing in child custody evaluations must differentiate between competent and incompetent use of psychological tests, then persuasively demonstrate the existence of a better methodology for obtaining the information traditionally garnered through psychological testing. In their article, *Individual Adult Psychometric Testing and Child Custody Evaluations: If the Shoe Doesn't Fit, Don't Wear It*, Benjamin Garber and Robert Simon, two psychologists, accomplish neither of these objectives. And in lieu of distinguishing between appropriate and inappropriate uses of tests or proffering a better substitute, they promote a less valid and reliable forensic child custody evaluation criterion—clinical judgment.³

We use the phrase, “*Anti-Testing, Anti-Individual, Pro-Family System Woozle*,” to capture the attempt to place a moratorium on psychological testing, shift attention away from consideration of individual dynamics in custody matters, and substitute a relatively undefined and untested “systems-level” approach to child custody evaluations and best interests

¹ See James N. Bow, James R. Flens & Jonathan W. Gould, *MMPI-2 and MCMI-III in Forensic Evaluations: A Survey of Psychologists*, 10 J. FORENSIC PSYCHOL. PRAC. 37 (2010); Robert E. Emery, Randy K. Otto & William T. O'Donohue, *A Critical Assessment of Child Custody Evaluations*, 6(1) PSYCHOL. SCI. IN THE PUB. INTEREST 1 (2005); Jonathan W. Gould, David A. Martindale & James R. Flens, *Use of Psychological Tests in Child Custody Evaluations*, in THE SCIENTIFIC BASIS OF CHILD CUSTODY DECISIONS __ (Robert M. Galatzer-Levy, Louis Kraus & Jeanne Galatzer-Levy eds. 2009); Randy K. Otto, John F. Eden & Elizabeth H. Barcus, *The Use of Psychological Testing in Child Custody Evaluations*, 38 FAM. CONCILIATION CT. REV. 312 (2000).

² MARC J. ACKERMAN, ANDREW KANE, JONATHAN GOULD, & MILFRED DALE, *PSYCHOLOGICAL EXPERTS IN DIVORCE ACTIONS* (6th ed. 2015).

³ Daniel Shuman & Bruce Sales, *The Admissibility of Expert Testimony Based upon Clinical Judgment and Scientific Research*, 4 PSYCHOL., PUB. POL'Y & L. 1226 (1998).

determinations.⁴ “Woozling” is the process of misrepresenting data⁵ and “woozles” are mistaken beliefs the misrepresentations create.⁶ Under the guise of science, scholar advocates can create controversy by justifying their positions without providing a complete representation of scientific and professional literature, by reporting only supporting literature, and by omitting explanations addressing both plausible alternative hypotheses and reasonable limitations of their opinions.⁷ We believe this has occurred in the *Shoe Doesn't Fit* article and collectively write with the message: Don't buy the “*Anti-Testing, Anti-Individual, Pro-Family Systems Woozle!*”

Few topics in child custody are without controversy or disagreement. Despite widespread use of mental health experts by family courts, challenges to mental health testimony,⁸ child custody evaluations,⁹ and psychological testing¹⁰ are not new. Some critics of mental health testimony and custody evaluations call for a moratorium on ultimate opinion testimony by mental health experts, arguing that there is a lack of scientific foundation for ultimate issue testimony.¹¹ Many theorists have questioned forensic use of a category of psychological tests, particularly projective psychological tests, because of concerns about psychometric

⁴ Benjamin Garber & Robert A. Simon, *Individual Adult Psychometric Testing and Child Custody Evaluations: If the Shoe Doesn't Fit, Don't Wear It*, 30 J. AM. ACAD. MATRIM. LAW. 325 (2018).

⁵ Robert Gelles, *The Politics of Research: The Use, Abuse, and Misuse of Social Science Data? The Cases of Intimate Partner Violence*, 45 FAM. CT. REV. 42 (2007).

⁶ Linda Nielson, *Pop Goes the Woozle: Being Misled by Research on Child Custody and Parenting Plans*, 56 J. DIVORCE & REMARRIAGE 595 (2015); See also Linda Nielson, *Woozles: Their Role in Custody Law Reform, Parenting Plans and Family Court*, 20 PSYCHOL., PUB. POL'Y & L. 164 (2014).

⁷ Robert E. Emery, Amy Holtzworth-Munroe, Janet R. Johnston, JoAnne L. Pedro-Carroll, Marsha Kline Pruett, Michael Saini & Irwin Sandler, “*Bending*” Evidence for a Cause: *Scholar-Advocacy Bias in Family Law*, 54 FAM. CT. REV. 134 (2016).

⁸ See Timothy M. Tippins & Jeffrey P. Wittmann, *Empirical and Ethical Problems with Custody Recommendations: A Call for Clinical Humility and Judicial Vigilance*, 43 FAM. CT. REV. 193 (2005).

⁹ Emery et al., *supra* note 1.

¹⁰ Steven K. Erickson, Scott O. Lilienfeld & Michael J. Vitacco, *A Critical Examination of the Suitability and Limitations of Psychological Tests in Family Court*, 45 FAM. CT. REV. 157 (2007).

¹¹ See Tippins & Wittmann, *supra* note 8, at 198.

reliability and validity.¹² Other commentators encourage use of tests with validity indices while discouraging those without such safeguards.¹³ Occasionally, specific psychological tests have been singled out for criticism (e.g., MCMI-III).¹⁴ Few go so far as the *Shoe Doesn't Fit* article, the authors of which propose a moratorium on psychological testing without, from our perspective, providing sufficient scientific support for their position.¹⁵ Even worse, they propose no method by which the data currently obtained through psychological testing might be obtained and similarly analyzed if a moratorium were to be placed on psychological testing. In part I of this article, we evaluate the position against the admission of psychological tests in family law cases. In part II we explain the distinction between psychological testing and psychological assessment and discuss admissibility issues as well as the importance of assessing parents' psychological functioning in custody evaluations. In part III, the need for empirically based evaluations is presented. In part IV, we review the risk of omitting psychological testing in custody evaluations. Finally, we conclude with a review of the problems with the Garber and Simon position and why we believe psychological testing can be an important part of a custody evaluation.

I. The “Anti-Testing” Position

Three types of *anti-testing* statements are scattered throughout the *Shoe Doesn't Fit* article. Identifying these three types of arguments, the manner in which they are made, and the scientific data relevant to each argument helps identify the strengths and weaknesses of the anti-testing position as well as where we collectively disagree.

¹² Erickson, Lilienfeld & Vitacco, *supra* note 10, at 158 (claiming that “most projective tests have substantial flaws that raise serious questions about their usefulness in these cases.”).

¹³ Richard Rogers & S.D. Bender, *Evaluation of Malingering and Related Response Styles*, in 11 COMPREHENSIVE HANDBOOK OF PSYCHOLOGY: FORENSIC PSYCHOLOGY 517 (Randy K. Otto ed., 2d ed. 2013).

¹⁴ Frank J. Dyer, *Forensic Applications of the MCMI-III in Light of Current Controversies*, in NEW DIRECTIONS IN INTERPRETING THE MILLON CLINICAL MULTIAXIAL INVENTORY-III 201 (Robert J. Craig, ed., 2015).

¹⁵ Garber & Simon, *supra* note 4, at 340.

The first anti-testing argument claims that little is known about custody litigants and that the psychological tests used to evaluate them lack normative data, making it impossible to accurately interpret test responses and scores.¹⁶ These claims are simply not true. This argument ignores the voluminous and ever-expanding data filling books, journals, and conferences about children of divorce, divorcing and separating parents, treatment interventions for these people, and, the equally robust literature on how to evaluate and help these children, parents, and families.¹⁷ This claim also ignores the reality that many of the psychological tests have normative data developed with custody litigants.¹⁸ For example, there are normative data sets for the Minnesota Multiphasic Personality Inventory – 2 (MMPI-2), Minnesota Multiphasic Personality Inventory – 2 – Restructured Form (MMPI-2-RF), Rorschach-Performance Assessment System (R-PAS), Personality Assessment Inventory (PAI), and several others.¹⁹

¹⁶ *Id.* at 329-330 (stating, “Experience suggests that they come from all socio-economic strata, all racial, religious, language and cultural groups, and span the full range of ages, IQs, sizes, shapes and colors. . . . Unless and until this population is better understood and instruments are developed that represent its normative thinking, feeling and behavior, it is simply misleading to make statements about a custody litigant by comparison to the responses of entirely distinct normative groups”).

¹⁷ Sol R. Rappaport, *Deconstructing the Impact of Divorce on Children*, 43 *FAM. L.Q.* 353 (2013). See also Lyn Greenberg, Lynda Doi Fick, & Robert Schnider, *Catching Them Before Too Much Damage Is Done: Early Intervention with Resistance-Refusal Dynamics*, 54 *FAM. CT. REV.* 548 (2016)

¹⁸ Kay Bathurst, Allen W. Gottfried & Adele E. Gottfried, *Normative Data for the MMPI-2 in Child Custody Litigation*, 9 *PSYCHOL. ASSESSMENT* 205 (1997); Robert E. Erard, Jacqueline S. Singer & Donald J. Viglione, *Multimethod Assessment with the Rorschach in Child Custody Cases*, in *THE RORSCHACH IN MULTIMETHOD FORENSIC ASSESSMENT: CONCEPTUAL FOUNDATIONS AND PRACTICAL ASSESSMENTS* 210 (Robert E. Erard & F. Barton Evans, eds., 2017).

¹⁹ Dan Hynan, *Use of the Personality Assessment Inventory in Child Custody Evaluation*, 5 *OPEN ACCESS J. FORENSIC PSYCHOL.* 122 (2013) (data for PAI); See also Elizabeth Archer, Leigh Hagan, Janelle Mason, Richard Handel & Robert Archer, *MMPI-2-RF Characteristics of Custody Evaluation Litigants* 19(1) *ASSESSMENT* 16 (2012) (data for MMPI-2-RF); Bathurst, Gottfried & Gottfried, *supra* note 18, at 205 (MMPI-2 data); Robert Erard, Jacqueline Singer, & Donald Viglione, *The Rorschach in Multimethod Custody Evalua-*

Child custody evaluations assisting attorneys and courts with custody decisions are the most complex and difficult type of forensic evaluations.²⁰ Competent psychological testing in these cases combines nomothetic evaluation with ideographic evaluation.²¹ Nomothetic interpretation examines test scores via comparison with a normative group, while ideographic interpretation involves understanding the test data within the context of historical data, presenting problems, observations, interview results, and information from third parties so as to examine competing hypotheses.²² *Best interests* determinations require decision-making specific to each individual child within his or her family. Ideographic interpretations in child custody evaluations include integration of several nomothetic interpretations of independent test scores with collateral information, interview data, collateral source and record data, and clinical judgment.²³

For example, a nomothetic interpretation from MMPI-2 data may indicate that a parent's scores on measures of paranoid functioning are similar to those characterized as being mistrusting, delusional, and displaying other behaviors associated with paranoid functioning. At this point in the analysis, there is nothing in the test results that informs the evaluator that this particular parent is paranoid or that the parent's possible paranoid functioning affects parenting, parent-to-parent communication, or the child's sense of safety, security, and predictability with that parent.

The evaluator's task is to explore whether any independent sources of data indicate that this parent's behavior is suggestive of paranoid functioning and, if so, in what areas relevant to parenting. Then, the evaluator must investigate how such paranoid behavior functionally affects various dimensions of parenting or co-parenting. Too often, evaluators report a score of clinical concern, such as a high score on a scale measuring para-

tions, in THE RORSCHACH IN MULTIMETHOD FORENSIC ASSESSMENT, *supra* note 18, at 210 (Rorschach data).

²⁰ Otto, Eden & Barcus, *supra* note 1.

²¹ George J. Meyer et al., *Psychological Testing and Psychological Assessment: A Review of Evidence and Issues*, in *METHODOLOGICAL ISSUES & STRATEGIES IN CLINICAL RESEARCH* 265 (Alan E. Kazdin ed. 3d ed. 2003).

²² LORRAINE DITTRICH EYDE ET AL., *RESPONSIBLE TEST USE: CASE STUDIES FOR ASSESSING HUMAN BEHAVIOR* (1993).

²³ See Gould, Martindale & Flens, *supra* note 1, at 85.

noid functioning but fail to investigate further whether such apparent paranoid behavior functionally affects various dimensions of parenting or co-parenting.

A second type of anti-testing argument says that because incompetent testing occurs, all testing should be banned. Such examples include the use of psychological tests for purposes other than those for which they were intended. Measures of personality constructs are not measures of parenting. Competent evaluators do not “blindly rely exclusively on computer-generated test scoring and interpretations.”²⁴ In fact, several peer-reviewed articles have directly addressed the substantial professional practice and ethical concerns about relying exclusively on computer-generated interpretations.²⁵

Psychological tests cannot “generate dispositive data regarding the psychological characteristics of parents as they pertain to parenting the children in the family dynamic”²⁶ nor can any psychological test dictate an outcome “against a best interests of the child criteria.”²⁷ Though the authors of this article acknowledge that inappropriate use of psychological test results occur in far too many cases,²⁸ this does not form the foundation for an argument that no testing should be done. The answer lies in requiring competence through training, supervision, and experience. The answer is not throwing the testing baby out with the bathwater.

Another argument against all testing relies on a misapprehension about the appropriate use and reporting of psychological testing in general. For example, test-retest reliability is a measure of the test’s consistency in measuring what it purports to

²⁴ Garber & Simon, *supra* note 4, at 327.

²⁵ Gould et al., *supra* note 1, at 85.

²⁶ PHILIP M. STAHL & ROBERT A. SIMON, FORENSIC PSYCHOLOGY CONSULTATION IN CHILD CUSTODY LITIGATION: A HANDBOOK FOR WORK PRODUCT REVIEW, CASE PREPARATION, AND EXPERT TESTIMONY 134 (2013) (stating that “[t]he smart evaluator and the state-of-the-art family law attorney recognize that psychological test data should only be used to generate hypotheses or identify areas of inquiry.”).

²⁷ Garber & Simon, *supra* note 4, at 336.

²⁸ James R. Flens, *The Responsible Use of Psychological Testing in Child Custody Evaluations: Selection of Tests*, 2 J. CHILD CUSTODY 3, 17 (2005) (regarding concerns with use of computer-generated test interpretation). *See also* Otto, Eden & Barcus, *supra* note 1 (regarding the proper use of psychological testing in custody evaluations).

412 *Journal of the American Academy of Matrimonial Lawyers*

measure across administrations absent a potentially transformative event.²⁹ The following statement by Garber and Simon has no bearing on test-retest reliability: “[I]t remains to be seen whether a properly normed test could demonstrate test-retest reliability across the period before, during and after the close of litigation.”³⁰ When a group of seventh grade campers come back after a summer away from home and their parents measure their height, and the parents encounter noticeably higher numbers on the tape measure than they saw before their children left for the summer, they don’t question the reliability of their scales. They recognize that the thing being measured has changed. Thus, this statement would provide support for elimination of testing only if we were to assume that the personal and environmental conditions affecting the parent at the time of first testing would be relatively the same at the second testing. Parental attitudes, emotions, and behaviors change over time, as does the litigation context within which the evaluation process occurs. Some of these changes increase distress and negatively affect parental functioning, while other changes diminish stress and improve parental functioning.

Similarly, we do not believe reporting basic statistics (e.g., “means” and “percentiles”) is “intimidating” or “risks misleading consumers and harming the children whom we intend to serve.”³¹ A review of the two articles cited as authority for this statement in the *Shoe Doesn’t Fit* reveals that one of the articles makes no such claim³² and the other addresses violence risk appraisal assessments, not testing in child custody matters.³³

We acknowledge that the risk exists that some attorneys and judges place greater weight on psychological test results when numbers are included in the custody advisory report. Numbers often create an aura of objectivity and precision, leading readers of evaluators’ reports to view test data as more meaningful than

²⁹ See ANNE ANASTASI & SUSANA URBINA, *PSYCHOLOGICAL TESTING* (7th ed. 1997).

³⁰ Garber & Simon, *supra* note 4, at 330.

³¹ *Id.*

³² Erickson, Lilienfeld & Vitacco, *supra* note 10, at 157.

³³ Stephen D. Hart, Christine Michie & David J. Cooke, *Precision of Actuarial Risk Assessment Instruments: Evaluating the ‘Margins of Error’ of Group v. Individual Predictions of Violence*, 190 *BRIT. J. PSYCHIATRY* 0-65 (May 2007).

other types of data in the report. Evaluators, however, can mitigate possible misunderstanding of the meaning of the scores by explaining their meaning and, in particular, framing the explanation within the published information pertaining to context-specific normative data for male and female custody litigants.

Data generated from psychological testing are not intended for indiscriminate use. We disagree with the *Shoe Doesn't Fit* conclusion that "individual adult psychometrics have little or no place in the process of evaluating family dynamics."³⁴ The usefulness of an evaluator's report can be significantly enhanced when psychological tests have been selected using generally agreed upon selection criteria,³⁵ when those tests have been competently administered and scored, and when the resulting data have been properly interpreted and integrated with other data. Properly interpreted test data often furnish information not available through other sources. The argument to completely exclude psychological testing and replace it with an amorphous realm of clinical judgment testimony ignores the wisdom of the legal system's requirements for admissibility of evidence.

Experts offering opinions, whether in court or in peer-reviewed articles, have a responsibility to accurately present the scientific and professional knowledge of the discipline, and to articulate the limitations of their opinions. Forensic psychologists also have an ethical obligation to consider reasonable alternative explanations. Guidelines governing child custody evaluations require the consideration of reasonable alternative explanations of data. In short, the ethical rules of the profession warrant consideration of appropriate available information—such as psychological tests—as well as consideration of the limitations of those instruments. Introduction of evidence that comports with the three "c"s of the scientific method—cumulative, comprehensive, and converging evidence—will promote better reasoned and better supported custody decisions.³⁶

In the next section, we distinguish between individual psychological test administration and psychological assessment, a

³⁴ Garber & Simon, *supra* note 4, at 326.

³⁵ Kirk Heilbrun, *Child Custody Evaluation: Critically Assessing Mental Health Experts and Psychological Tests*, 29 FAM. L.Q. 63 (1995).

³⁶ See, e.g., Nancy Levit, *Listening to Tribal Legends: An Essay on Law and the Scientific Method*, 58 FORDHAM L. REV. 263, 267 (1989).

distinction that Garber and Simon fail to make. We then provide a review of the relevance of standards in the use of psychological tests in child custody evaluations. We then will address issues related to admissibility and argue that the exclusion of psychological testing may decrease, not increase, the admissibility of child custody evaluations. We end this section addressing what we believe would be the likely negative impacts on the field of forensic psychology if courts were to exclude psychological tests and assessments.

II. Understanding Psychological Testing/ Assessment

Psychological testing is a relatively straightforward process. A particular test is administered to obtain a specific score. A descriptive meaning can be applied to the score based on data obtained from a normative group.³⁷ This has been called *nomothetic* evaluation because the interpretation of the score is based entirely upon comparison with the normative group. Interpretation does not include an examination of the individual's score within the context of the person's life. In child custody evaluations, most interpretations of individual psychological tests begin as nomothetic interpretations because we are comparing scores from each parent or child with scores from a normative group.³⁸

Psychological assessment includes, but is not limited to, psychological testing.³⁹ In *psychological assessment*, the evaluator takes a variety of scores generally obtained from multiple test methods and considers "the data in the context of history, referral information, and observed behavior to understand the person being evaluated, to answer referral questions, and then to communicate findings."⁴⁰ This has been called *ideographic* evaluation because the assessment task is to understand the psychological test score within the context of the person's life. The assessment task is to use test-derived sources of information in combination with historical data, presenting complaints, obser-

³⁷ Meyer et al., *supra* note 21, at 143.

³⁸ Gould, Martindale & Flens, *supra* note 1, at 85.

³⁹ Meyer et al., *supra* note 21, at 143.

⁴⁰ *Id.* at 265.

vations, interview results, and information from third parties in order to examine competing hypotheses.⁴¹ Ideographic assessment requires a high degree of skill and sophistication.⁴²

In the context of child custody evaluations, the ideographic interpretation of psychological test data entails the integration of information from multiple sources that describe the parents, the children, and parent-child interactions. Ideographic interpretations found in child custody evaluations include integration of several nomothetic interpretations of independent test scores with collateral information, interview data, collateral record data, and clinical judgment.⁴³

A. Choosing Assessment Instruments

The question to be asked (and correctly answered) is: “Do the psychological tests you have elected to utilize reliably measure *functional* abilities which bear directly upon the matter before the court?” It is reasonably expected that evaluators will obtain, review, and critically examine the documentation concerning assessment instruments under consideration and they will select only those instruments whose reliability and validity are acceptable when used for the purpose intended.⁴⁴ In a child custody evaluation, understanding a parent’s psychological functioning is not solely about understanding if there is psychopathology, but also gaining an understanding of a parent’s ability to regulate and modulate emotions, coping capacities under stress, capacity for empathy, among other attributes, all of which can be garnered with the assistance of psychological testing.

In order to critically examine the role of testing in custody evaluations, we must consider what type of information is sought, how tests may (or may not) assist in gathering pertinent information, and the ways in which tests of every type are deficient. The search for pertinent information is impeded when theorists and researchers look in the wrong places. The proper and limited use of psychological testing in child custody indicates that it has tremendous value to judges making custody determinations.

⁴¹ EYDE ET AL., *supra* note 22; KIRK HEILBRUN, PRINCIPLES OF FORENSIC MENTAL HEALTH ASSESSMENT (2001).

⁴² Meyer et al., *supra* note 21, at 144.

⁴³ Gould, Martindale & Flens, *supra* note 1, at 86.

⁴⁴ *Id.*

The authors of the *Standards for Educational and Psychological Testing* have observed: “The greater the potential impact on test takers, for good or ill, the greater the need to identify and satisfy the relevant standards.”⁴⁵ Jonathan Gould, David Martindale, and James Flens extended that obligation further, emphasizing that the greater the potential impact on consumers of psychological services, for good or ill, the greater the need to adhere to established standards, to be responsive to applicable guidelines, and, to endeavor to identify and subsequently utilize the best methodology possible.⁴⁶

B. *Daubert, Frye*, and Admissibility Issues

In 1971, a case came before the U.S. Supreme Court which had nothing whatsoever to do with custody work but the reverberations of which have been dramatically felt by evaluators. The decision in *Griggs v. Duke Power Co.*, called the attention of custody evaluators to what have come to be known as *functional abilities*.⁴⁷ In *Griggs*, the Court ruled that “testing procedures must be demonstrably reasonable measures of (or predictors of) job performance.”⁴⁸ Psychologists’ endeavors to assess the characteristics that bear directly upon parenting are more likely to meet with success if they conceptualize parenting as a job and focus attention on those attributes, behaviors, attitudes, and skills that are reliably related to the demands of the job. Though no polls have been taken on the matter, it is our impression there is agreement among experienced evaluators that it is advisable to conceptualize parenting as a skill set and to choose instruments and techniques that will assess those skills.⁴⁹ We argue, as will be described below, that the proper use of psychological testing can assist an evaluator in understanding the functional abilities of parenting.

⁴⁵ AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, AMERICAN PSYCHOLOGICAL ASSOCIATION, NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION, STANDARDS FOR EDUCATIONAL AND PSYCHOLOGICAL TESTING 112 (1999).

⁴⁶ Gould, Martindale & Flens, *supra* note 1, at 85.

⁴⁷ *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971).

⁴⁸ David A. Martindale & Jonathan W. Gould, *The Forensic Model*, 1 J. CHILD CUSTODY 1 (2004).

⁴⁹ Gould, Martindale & Flens, *supra* note 1, at 85.

States that have adopted *Daubert v. Merrell Dow Pharmaceuticals, Inc.*⁵⁰ factors in addressing the admissibility of scientific evidence base their analyses on the expert's qualifications and the reliability and relevance of the proffered evidence. Among the questions often asked in a *Daubert* challenge include whether the expert's opinion is based upon incomplete data, insufficient data, consideration of reasonable alternative explanations of data, and the degree to which proffered opinions are related to the information upon which they are based.⁵¹

Expert opinions based upon a model of child custody assessment examining only a family system's level of analysis may not pass a *Daubert* challenge. There are many other levels of analysis that may be important to examine, which may include, but are not limited to, individual parent factors, individual child factors, parent-child relationship factors, extended family factors, and community factors.

For example, there are individual factors that may be critically important to assess in a child custody evaluation. There is considerable empirical research examining the individual level of analysis pertaining to parenting behavior correlates associated with personality disorders. These empirical studies have found that different personality disorders are often associated with predictable problematic parenting behaviors⁵² and risk for child maltreatment.⁵³ Other studies have found that individuals with different personality disorders who display particular predictable problematic parenting behaviors affect children in predictable ways.⁵⁴

⁵⁰ 509 U.S. 579 (1993).

⁵¹ See, e.g., David M. Flores et. al., *Examining the Effects of the Daubert Trilogy on Expert Evidence Practices in Federal Civil Court: An Empirical Analysis*, 34 S. ILL. U. L.J. 533 (2010).

⁵² Sarah Laulik, Shihning Chou, Kevin D. Browne & Jayne Allam, *The Link Between Personality Disorder and Parenting Behaviors: A Systematic Review*, 18 AGGRESSION & VIOLENT BEHAV. 644 (2013).

⁵³ Sarah Laulik, Jayne Allam & Kevin D. Browne, *Maternal Borderline Personality Disorder and Risk of Child Maltreatment*, 25 CHILD ABUSE REV. 300 (2016).

⁵⁴ Julie Eyden, Catherine Winsper, Dieter Wolke, Matthew Richard Broome & Fiona MacCallum, *A Systematic Review of the Parenting and Outcomes Experienced by Offspring of Mothers with Borderline Personality Pathol-*

Psychological testing may provide valuable information about the degree to which a parent's measured personality characteristics may be similar to those of individuals diagnosed with specific personality disorders. Excluding information about individual personality factors and the relationship between the behavioral correlates associated with deficient parenting behaviors may fail a *Daubert* challenge precisely because of the potential value inherent in the data generated by testing aimed at an individual level of analysis.

C. The Challenge that Psychological Tests Are not Normed with Litigants

Garber and Simon posit in their abstract that because psychological tests are not normed with litigating parents, the tests cannot provide valid data. Based in part on this statement, they conclude that evaluators consider not using psychological testing. This is a straw man argument. First, the statement that psychological tests are not normed with litigating parents is inaccurate for several of the tests. According to the Standards for Educational and Psychological Testing, which is co-published by the American Psychological Association, American Educational Research Association, and the National Council on Measurement in Education, norms are “[s]tatistics or tabular data that summarize the distribution or frequency of test scores for one or more specified groups, such as test takers of various ages or grades, usually designed to represent some larger population, referred to as the *reference population*.”⁵⁵ Furthermore, the above standards define local norms as well. Local norms are “[n]orms by which test scores are referred to a specific, limited reference population of particular interest to the test user (e.g., population of a locale, organization, or institution). Local norms are not intended to be representative of populations beyond that limited setting.”⁵⁶ For many of the most commonly used tests they describe in their arti-

ogy: *Potential Mechanisms and Clinical Implications*, 47 *CLINICAL PSYCHOL. REV.* 85 (2016).

⁵⁵ AMERICAN EDUCATIONAL RESEARCH ASSOCIATION ET AL., *supra* note 45, at 112 (emphasis added).

⁵⁶ *Id.* at 220.

cle, there are norms for custody litigants.⁵⁷ At the same time, we acknowledge there are not specific validity studies on these norms, and recognize this as a limitation of the tests. However, Garber and Simon fail to point out that for the most part, the normative reference samples of custody litigants for the various tests show custody litigants score similarly to the general population on most scales.⁵⁸ The data also show custody litigants tend to respond more defensively than most people.⁵⁹ However, this is not surprising, nor does it mean testing does not have value. In fact, there are several other groups that score defensively on tests, such as police officers in their pre-employment screenings. This does not negate the use of testing. Furthermore, psychological testing is commonly used with personal injury litigants in a similar manner to custody litigants, in that evaluators can often compare a personal injury litigant's score against the scores of other personal injury litigants. Furthermore, the fact that custody litigants are a specific group of adults, and are not identical to the population upon which the test was initially developed, does not mean there is no value in administering psychological tests. However, it does mean evaluators need to acknowledge reasonable limitations of the proffered interpretations. Finally, one of the main reasons why the argument that tests should be discarded because not all are normed on litigants fails is because proper test interpretation does not mean a score is interpreted identically for every evaluatee. As will be described later, psychological test results provide hypotheses about a person, rather than directly state a person has a specific attribute or pathology.

D. *The Importance of Understanding Parents' Psychological Functioning*

Garber and Simon also state the question to be addressed in a child custody evaluation (CCE) is not about the individual adult functioning, and therefore, testing is unnecessary. While

⁵⁷ Dan Hynan, *Use of the Personality Assessment Inventory in Child Custody Evaluation*, 5 OPEN ACCESS J. FORENSIC PSYCHOL. 122 (2013) (data for PAI); See also Archer et al., *supra* note 19 (data for MMPI-2-RF); Bathurst, Gottfried & Gottfried, *supra* note 18, at 205 (MMPI-2 data); Erard et al., *supra* note 19, at 1 (Rorschach data).

⁵⁸ Bathurst, Gottfried & Gottfried, *supra* note 18, at 205.

⁵⁹ *Id.*

the focus of a CCE is on the psychological best interests of a child, evaluators and the court need to understand what each parent brings to the child(ren), which includes their psychological functioning, and its impact on parenting. Garber and Simon appear to make an *a priori* assumption that there is little value in a proper assessment of individual psychological factors in family conflict and parenting behavior because all the relevant information needed to answer questions about custodial access may be found in direct observations and third party information about parent-child interactions. This is an unsupported thesis that actually proves too much—if it were true, expert testimony would never be needed in court.

The American Psychological Association's Guidelines for Child Custody Evaluations in Divorce Proceedings identify the purpose of a child custody evaluation as to assist in determining the psychological best interests of the child.⁶⁰ "When conducting child custody evaluations, psychologists are expected to focus on factors that pertain specifically to the psychological best interests of the child, because the court will draw upon these considerations in order to reach its own conclusions and render a decision."⁶¹ Psychologists are "encouraged to weigh and incorporate such overlapping factors as family dynamics and interactions; cultural and environmental variables; relevant challenges and aptitudes for all examined parties; and the child's educational, physical, and psychological needs."⁶²

Child Custody Guideline 3 reads "The evaluation focuses upon parenting attributes, the child's psychological needs, and the resulting fit."⁶³ Among the relevant commentary provided in the guidelines is the "most useful and influential evaluations focus upon skills, deficits, values, and tendencies relevant to parenting attributes and a child's psychological needs. Comparatively little weight is afforded to evaluations that offer a general personality assessment without attempting to place results in the appropriate context."⁶⁴ We agree that comparatively little

⁶⁰ American Psychological Association, *Guidelines for Child Custody Evaluations in Family Law Proceedings*, 65 AM. PSYCHOL. 863 (2010)

⁶¹ *Id.* at 864.

⁶² *Id.*

⁶³ *Id.*

⁶⁴ *Id.*

weight should be afforded to general personality assessment without demonstrating how a parent's particular psychological make-up affects parenting competencies, parent-child relationship and interactions, parent-to-parent communication, and other similarly relevant factors. We disagree that the data obtained through proper testing provides little, if any, useful information about any of these factors. In fact, many of the writers who have co-authored this article have written and spoken extensively about the usefulness of psychological testing and its limitations for more than twenty years.⁶⁵ While we acknowledge the importance of attending to family system level of analysis, we also believe it is critical to examine the contribution of parental psychopathology and dysfunction since it may affect parenting behavior, the parent-child relationship, and parent-to-parent cooperation and communication.

There is a wealth of research on the impact of various parental mental health problems on children, such as parental Depression, Borderline Personality Disorder, and Anxiety Disorders, among others.⁶⁶ Furthermore, parental traits such as warmth and empathy can impact children. Understanding a parent's mental health and their ability to provide warm and empathic care is a vital part of conducting a CCE. We are not suggesting psychological testing must be done in every CCE. Where we disagree with Garber and Simon is that to suggest that psychological testing is not needed because the focus of the evaluation is not on the adult's individual functioning is seriously flawed. If this underlying assumption were accepted, this could inappropriately lead evaluators to not ask about a parent's own mental health, or could lead to evaluators not talking to collateral sources, such as a parent's therapists, to better understand the parent's psychological functioning. While we are not advocating for testing children, we are advocating child custody evaluators, guardians ad

⁶⁵ See JONATHAN W. GOULD, CONDUCTING SCIENTIFICALLY CRAFTED CHILD CUSTODY EVALUATIONS (2d ed. 2006); Flens, *supra* note 28. Sol Rappaport, Psychological Testing and Cross-examination, Oklahoma Bar Ass'n – Tulsa, Oklahoma (2011).

⁶⁶ Jenny Macfie, *Development in Children and Adolescents Whose Mothers Have Borderline Personality Disorder*, 3 CHILD DEV. PERSPECTIVES 66 (2009); See also Carolyn Zahn-Waxler, Sunita Duggal & Reut Gruber, *Parental Psychopathology*, in 4 HANDBOOK OF PARENTING: SOCIAL CONDITIONS AND APPLIED PARENTING 295 (Marc Bornstein ed., 2d ed. 2002).

litem, the judge, and other family law professionals involved in a case need to understand the child's functioning as well as the parents', and that the sole focus cannot be just on the parent-child relationship. Child custody evaluators, attorneys, and judges cannot fully understand a family system without understanding the parts that make up the system. We believe that the Garber-Simon focus on parent-child fit as the *sine qua non* of evaluation focus is an example of painting by the numbers. It is a one-size-fits-all solution to complex family system challenges that often have outlived noble attempts at resolution through alternative dispute resolution.

The *Shoe Doesn't Fit* article fails to consider that many states have statutory requirements in child custody evaluations to address the psychological functioning of each individual. For example, in the California Rules of Court, Rule 5.220(C)(5) states that the custody evaluation should including an assessment of "psychological and social functioning."⁶⁷ In Illinois, section 602.7 of the Illinois Marriage and Dissolution of Marriage Act states that in making decisions regarding parenting time the "mental and physical health of all individuals involved" must be considered.⁶⁸ Thus, the article, while proposing a change in focus of custody evaluations, would encourage evaluators to bypass statutory requirements.

The elimination of psychological testing from child custody evaluations removes a procedure that may be useful in generating plausible alternative hypotheses to consider with regard to factors operating within a family system. Garber and Simon appear to disregard the potential contribution of an individual's psychological dysfunction as a problematic source of family conflict. Worse, in our view, is they appear to rely on their clinical judgment drawn from observational data about an individual's potential psychopathology and its effect on parenting behavior, parent-child interaction, and parent-to-parent cooperation and communication.

Garber and Simon compare the use of medical tests, such as Functional Magnetic Resonance Imaging (fMRI) tests and X-rays, to psychological tests, and unsurprisingly find the latter to

⁶⁷ CAL. RULES CT., R. 5.220(e)(2)(C)(v).

⁶⁸ 750 ILL. COMP. STAT. § 5/620.7(7) (2017).

be wanting. In doing so, they construct an argument in which their conclusion—to eliminate testing—does not follow from the premise. In these medical tests, the results show something within the body that cannot be seen. If someone has a fracture, an x-ray may show it. Physicians draw conclusions based on these tests. However, in psychological testing, evaluators do not look at a test score and draw a conclusion in a similar manner. Psychological testing provides hypotheses. For example, if a parent scores high on the Depression scale of the MMPI-2, an examiner should not conclude the parent is depressed. The correct way to interpret the above result is to develop a hypothesis that the parent may be depressed based on the test score. The evaluator then needs to look for other information which supports that conclusion as well as look for information which disconfirms it. We agree with Garber and Simon that there are evaluators who do not properly interpret tests. We are confident they would likely also state that there are x-ray technicians who misinterpret x-rays. The misinterpretation of a test does not mean the test is not valid or reliable, or cannot be used. It means the profession needs to improve its teaching in x-ray or psychological test interpretation. Overall, the comparison of psychological test interpretation to medical test interpretation is not a valid comparison, as there are two different interpretive strategies. In medical tests health care providers draw conclusions based on the test. In psychological testing, mental health professionals develop hypotheses.

In a similar vein, Garber and Simon state that evaluators use invalid means and alpha coefficients⁶⁹ in their reports. First, we do not know what they mean by invalid means. Second, and more importantly, among all of the authors of this reply, who have collectively reviewed more than one thousand custody evaluations, not one of us have seen alpha coefficients in a report. We agree evaluators should not use invalid means and alpha coefficients in their reports. However, we disagree that percentiles should not be used. In some situations there are valid reasons to express to the court a parent's score as a percentile. If this is done, the evaluator needs to explain what it means and the rele-

⁶⁹ Alpha coefficients measure the internal consistency of a scale. That is, whether the items that compose a particular scale on a test measure the same or similar ideas.

vance to the person referenced. The fact that some evaluators may misinterpret a test or report on the results in an incomplete manner, does not mean testing should not be done or has no value, although we agree the person who misuses a test should likely not be testing.

Garber and Simon state that custody evaluators should “eschew the use of individual adult psychometric tests in favor of cautiously and carefully crafted, systemically-informed, empirically-grounded, and child-centered *qualitative* observations inferences, and recommendations.”⁷⁰ First, the two are not mutually exclusive. Psychological testing can be a part of a cautiously and carefully crafted evaluation. Second, they reference using an empirically-grounded approach. Psychological testing is the most empirically based aspect of a custody evaluation. Science is “a set of systematic safeguards against confirmation bias, that is, the tendency to seek out evidence consistent with our hypotheses and to deny, dismiss, or distort evidence which runs counter to them.”⁷¹ Science requires a “willingness to accept facts even when they are opposed to wishes.”⁷² Others have talked about the purpose of science as “bending over backwards to prove ourselves wrong.”⁷³ Psychological testing can be used as a safeguard to protect against bias. Thus, the argument to discard a valuable source of information—to systematically ignore a slice of more comprehensive information—does not comport with either the principles of science or the criteria of rationality.

Garber and Simon go on to accurately describe one of the problems in using the Millon Clinical Multiaxial Inventory (MCMI) in a child custody evaluation. However, they failed to note that the manual for the MCMI states:

The use of MCMI in child custody evaluations is somewhat controversial. On the one hand, there is an understandable interest in determining each parent’s psychological health as one aspect of the ability to adequately discharge parental responsibilities. On the other hand, the MCMI-IV was developed and normed using a group that is far different from child custody litigants in terms of the circumstances surrounding the assessment. The clinician should become knowledgeable

⁷⁰ Garber & Simon, *supra* note 4, at 327.

⁷¹ Scott Lilienfeld, *Can Psychology Become a Science?*, 49 *PERSONALITY & INDIVIDUAL DIFFERENCES* 2 (2010).

⁷² B. F. SKINNER, *SCIENCE AND HUMAN BEHAVIOR* 12 (1953).

⁷³ *Id.*

about both sides of this controversy before deciding whether to include the MCMI-IV as part of a child custody evaluation.⁷⁴

Again, the issue is not if a test should or should not be used, but whether it is used appropriately in light of the research on the test. The misuse of a test in some instances does not mean a test should not be used in all instances. Garber and Simon also state it is “misleading to make statements about a custody litigant by comparison to the responses of other entirely distinct normative groups.”⁷⁵ Their proposition is an argument against the misuse of a test, not the use of it. Child custody evaluators have a responsibility to consider plausible alternative hypotheses. The Specialty Guidelines for Forensic Psychology Guideline 9.01 addresses the critical importance of considering plausible alternative hypotheses.⁷⁶

E. Psychological Tests and Data

Garber and Simon draw conclusions without data to support them. This may lead the reader to accept something as fact for which there is no support in the research. For example, Garber and Simon state that psychological test results are often given great weight. We are unaware of any research that would support such a broad statement. In fact, in the only study we are aware of which specifically addressed judicial decision making in custody disputes, there was little discussion about psychological testing by the judges in the study.⁷⁷ In fact, Anat Geva reported that “[s]everal judges reported expecting evaluators to conduct psychological tests as a part of the parental assessment but admitted to not being well informed regarding the characteristics of these tests.”⁷⁸ Geva went on to note that one judge said, “I just kind of flip over to see what the recommendation is. Where I need to look further is at the testing results and stuff in there . . .

⁷⁴ THEODORE MILLON, SETH GROSSMAN, & CARRIE MILLON, MILLON CLINICAL MULTIAxIAL INVENTORY, 9 (2015).

⁷⁵ Garber & Simon, *supra* note 4, at 329-330.

⁷⁶ American Psychol. Ass’n, *Specialty Guidelines in Forensic Psychology*, 68 AM. PSYCHOL. 7 (2013).

⁷⁷ Anat S. Geva, *Judicial Determination of Child Custody When a Parent Is Mentally Ill: A Little Bit of Law, a Little Bit of Pop Psychology, and a Little Bit of Comment Sense*, 16 U.C. DAVIS J. JUV. L. & POL’Y 1 (2012).

⁷⁸ *Id.* at 55.

I don't really focus on that very much."⁷⁹ Geva described in detail how judges made their decision in custody disputes; there is no evidence that psychological testing carried great weight.

Garber and Simon raise the issue of whether assessing people in the midst of a crisis may not capture their typical functioning. This is a valid concern. Part of an evaluator's job is to address this precise issue. Garber and Simon's concern is that a test administered in an evaluation may not show test-retest reliability before, during, and after the close of litigation. There are several flaws with this argument. While it is correct that custody litigants are in a difficult time of their life and the custody battle itself can induce or exacerbate mental health problems, this does not mean the court does not need information about their current functioning. Furthermore, this argument almost proves too much—it could be used to support the conclusion that one could not conduct a child custody evaluation during the crisis because one could not know for sure how the litigant will be six months later. We agree that evaluators can't entirely accurately predict a parent's future behavior, with or without testing. However, having an understanding of a parent's psychological functioning at the time of an evaluation is one piece of data the court often considers. In fact, most states' statutes specifically list factors that the court must consider in determining what is in the best interest of the child. In many states, it includes the mental health of the parties.⁸⁰ The evaluator's role is to collect a wealth of data, integrate it, and explain to the court what conclusions the evaluator reached, and how the evaluator reached the conclusions. The fact that a person may be under stress and duress during an evaluation does not preclude psychological testing from being of assistance.

Finally, as support for their argument, Garber and Simon cite the American Academy of Child and Adolescent Psychiatry Practice Parameters for Child Custody Evaluation. However, the article fails to mention that the same Practice Parameters state:

When the psychiatric health of a parent or child is a legitimate issue, the evaluator may request psychological testing of each parent to help

⁷⁹ *Id.*

⁸⁰ *See, e.g.,* MO. REV. STAT. § 452.375.2 (2016) (requiring consideration of “[t]he physical and mental health of all individuals involved”).

support an opinion and provide relevant data. This may add to the degree of certainty of the parenting assessment. Certain tests have been advanced as having specific utility in assessing variables specific to a custody evaluation. These include the Bricklin Perception of Relationships Test (Bricklin, 1995) and the Ackerman-Schoendort Scales for Parenting Evaluations of Custody (Ackerman, 1994). Use of these tests is controversial at present. Their role in a custody evaluation should be adjunctive and they should never take the place of a comprehensive evaluation.⁸¹

The article also fails to cite the American Academy of Child and Adolescent Psychiatry Practice Parameters for Child and Adolescent Forensic Evaluations, a document written fourteen years after the Practice Parameters the authors cite.⁸² The 2011 parameters state that “in child custody disputes when there are significant allegations and inconsistency in presentation, a comprehensive psychological battery, including tests such as the Minnesota Multiphasic Personality Inventory-2 Restructured Form and the Millon Clinical Multiaxial Inventory-III, can be helpful in further assessing personality, deceptiveness, and defensiveness in the parents.”⁸³

The *Shoe Doesn't Fit* article is an advocacy piece which fails to even acknowledge where guidelines cited in the article support the use of psychological testing.

III. The Need for Empirically Based Evaluations

Throughout their article, Garber and Simon reiterate the need for an empirically-driven examination. We agree with this. In fact, we believe psychological testing is the part of the evaluative process where the profession actually has the most empirical information, which is why we believe there is an appropriate place for psychological testing in custody evaluations. We agree with Garber and Simon, who cite Philip Stahl and Robert Simon for the proposition that psychological tests are not dispositive data regarding the characteristics of a parent.⁸⁴ As we have stated earlier, psychological test results provide hypotheses about

⁸¹ American Academy of Child and Adolescent Psychiatry Guidelines, *Practice Parameters for Child Custody Evaluation*, 36 J. AM. ACAD. CHILD & ADOLESCENT PSYCHIATRY 57S, 65S (1997).

⁸² *Id.*

⁸³ *Id.* at 1308.

⁸⁴ STAHL & SIMON, *supra* note 26, at 125.

a person's functioning, not dispositive data. At the same time, evaluators can conduct an item analysis to see what specific questions contributed to a scale being elevated. This can assist an evaluator, and ultimately the court, in understanding a parent's psychological functioning. Garber and Simon argue "the fact that testing generates quantitative data does not make it more objective, more valuable, or a better check or balance."⁸⁵ While we agree that test data is not inherently more valuable or a better check or balance, it is, in fact, more objective than an individual evaluator's clinical judgment (where $n = 1$). As the U.S. Supreme Court said in *Daubert*, "Expert testimony that is not grounded in methods and procedures acknowledged by scientists in the particular field of study amount to no more than subjective belief or unsupported speculation."⁸⁶ The use of psychological testing provides an evaluator with methods grounded in science, thus decreasing the likelihood of a single individual's subjective beliefs and unsupported speculations.

As support for their position, Garber and Simon state that the purpose of a CCE is not to generate psychiatric diagnoses. We wholeheartedly agree. The purpose of psychological testing in a custody evaluation is not to develop a diagnosis. Rather, it is to assist the court in understanding a parent's psychological functioning, and most importantly, how that functioning may, can, or does impact their ability to parent their child, their relationship with their child, and factors associated with cooperatively communicating with the other parent. Garber and Simon even cite some of the authors of this paper for the problems with psychological testing. While we are not suggesting psychological testing is a panacea, the citations they use against testing are more than ten years old, and were written prior to many of the tests being used having normative data for custody litigants. There is much more data available in the social science research on the use of specific tests in custody disputes than in the early 2000's. There is also growing data on the impact of parents' psychological functioning on children.⁸⁷

⁸⁵ Garber & Simon, *supra* note 4, at 335.

⁸⁶ *Daubert*, 509 U.S. at 589-90.

⁸⁷ See, e.g., Zahn-Waxler et al., *supra* note 66 (discussing the negative impact of various mental health problems in parents on children). See also ALISON GOPNIK, *THE GARDENER AND THE CARPENTER: WHAT THE NEW SCIENCE*

Our field has known of the significant threats to reliable expert testimony when such testimony is solely based upon clinical judgment. The concern about clinical judgment used in a forensic context is reflected in several recent articles. Daniel Shuman and Bruce Sales make the following points.

1. While some expert testimony by people with scientific degrees is derived from research the accuracy of which can be validated, much other expert testimony advances opinions derived from judgments in which accuracy rests on the experts' non-validated theories and skills;

2. These untested opinions are commonly referred to as "clinical" judgments . . . and are defined by their reliance on personal experience rather than on statistically analyzed data drawn from valid and reliable research;

OF CHILD DEVELOPMENT TELLS US ABOUT THE RELATIONSHIP BETWEEN PARENTS AND CHILDREN (2016); HANDBOOK OF ATTACHMENT: THEORY, RESEARCH, AND CLINICAL APPLICATIONS (Jude Cassidy & Phillip R. Shaver eds., 3d ed. 2016); Sunhye Bai, Rena L. Repetti, & Jaqueline B. Sperling, *Children's Expressions of Positive Emotion Are Sustained by Smiling, Touching, and Playing with Parents and Siblings: A Naturalistic Observational Study of Family Life*, 52 DEV. PSYCHOL. 88 (2016); Sophie Bridgers, Daphna Buchsbaum, Elizabeth Seiver, Thomas L. Griffiths & Alison Gopnik, *Children's Causal Inferences from Conflicting Testimony and Observations*, 52 DEV. PSYCHOL. 9 (2016); Rebecca L. Brock, Grazyna Kochanska & Lea J. Boldt, *Interplay Between Children's Biobehavioral Plasticity and Interparental Relationship in the Origins of Internalizing Problems*, J. FAM. PSYCHOL. 1 (2017); Kirby Deater-Deckard & Martha Ann Bell, *Maternal Executive Function, Heart Rate, and EEG Alpha Reactivity Interact in the Prediction of Harsh Parenting*, 31 J. FAM. PSYCHOL. 41 (2017); Steven Demby, *Commentary on Entrenched Postseparation Parenting Disputes: The Role of Interparental Hatred*, 55 FAM. CT. REV. 417 (2017); Eirini Flouri & Emily Midouhas, *Environmental Adversity and Children's Early Trajectories of Problem Behavior: The Role of Harsh Parental Discipline*, 31 J. FAM. PSYCHOL. 234 (2017); Erika Lunkenheimer, Nilam Ram, Elizabeth A. Skowron & Peifeng Yin, *Harsh Parenting, Child Behavior Problems, and the Dynamic Coupling of Parents' and Children's Positive Behaviors*, 31 J. FAM. PSYCHOL. 689 (2017); Nicola McConnell, Matt Barnard, Julie Taylor, *Caring Dads Safer Children: Families' Perspectives on an Intervention for Maltreating Fathers*, 7 PSYCHOL. VIOLENCE 406 (2017); Meike Slagt, Judith Semon Dubas, Maja Dekovic & Marcel A. G. van Aken, *Differences in Sensitivity to Parenting Depending on Child Temperament: a Meta-Analysis*, 142 PSYCHOL. BULL. 1068 (2016); Laura Sneidman & Amanda L. Woodward, *Are Child-Directed Interactions the Cradle of Social Learning?*, 142 PSYCHOL. BULL. 1 (2016).

3. Use of the term “clinical” refers to a method or approach of making judgments or decisions;

4. The growing literature on human judgment and decision making helps explain the inherent unreliability of clinical judgment and decision making;

5. Expert judgments that are clinically derived . . . are as susceptible to error as lay judgments and use strategies in arriving at decisions that contribute to the error rate;

6. Clinical judgments and opinions offered in court are just as flawed as any other clinical judgments;

7. To the extent that a scientist or practitioner is relying on personal experience and personal biases in drawing inferences that go beyond the data, he or she is engaging in clinical decision making, despite his or her scientific training.⁸⁸

From a legal perspective, these concerns about threats to reliable expert testimony have been addressed in court decisions that follow *Daubert* and its progeny. An expert is expected to rule out alternative possible causes with reasonable certainty⁸⁹ and the failure to consider plausible alternative hypotheses and/or causes renders expert opinion as little more than speculation.⁹⁰ Other courts have cited similar language in determining whether the expert has adequately accounted for obvious alternative explanations.⁹¹ Psychological testing provides data that is not solely clinical judgment. It provides information that can challenge an evaluator’s working hypothesis or provide support for it. It may also result in the generation of new hypotheses.

IV. The Hazards of Omitting Psychometric Test from Custody Evaluations

A. Discarding Valuable Evidence Could Open Judges’ Decisions to Reversal on Appeal

The danger of a court affirmatively discarding valuable evidence is that it would open the judge’s decision to attack for being based on incomplete information. First, contrary to Garber

⁸⁸ Shuman & Sales, *supra* note 3, at 1229.

⁸⁹ Merrell Dow Pharm., Inc. v. Hayner, 953 S.W.2d 706, 720 (Tex. 1997).

⁹⁰ E.I. Dupont de Nemours & Co. & Robinson, 923 S.W.2d 549, 559 (Tex. 1995).

⁹¹ State v. McGrady, 787 S.E.2d 1, 9 (N.C. 2016).

and Simon's assertion, there is normative data for custody litigants on many of the psychological tests used in custody evaluations. Second, psychological testing in custody disputes has repeatedly survived *Daubert* and other legal challenges. We are not sure why Garber and Simon think it will not. Third, the characteristics that tests are designed to measure are, in fact, central to understanding a parent's mental health, which can have a direct impact on parenting. As previously stated, many states' best interest statutes specifically say that the court must consider the mental health of each parent. We agree with Garber and Simon that using a model such as the one developed by Timothy Tippins and Jeffrey Wittman is a useful approach.⁹² However, we disagree with their application of that model. While we agree it is reasonable to generate hypotheses based on test scores, we disagree with their application of the remainder of the model. We believe the proper approach is not to draw conclusions from testing alone and then draw inferences about how a child's needs might best be served, but rather to use data from testing as one point of data in drawing conclusions.

Furthermore, assessing a parent's mental health is often required as part of a best interests of the child determination. The Uniform Marriage and Divorce Act of 1970 noted, "The court shall determine custody in accordance with the best interests of the child. The court shall consider all relevant factors including: . . . (5) the mental and physical health of all individuals involved."⁹³ All fifty states and the District of Columbia include considerations of all relevant factors in best interests determinations. Some states explicitly list the mental health of the parents as a factor and other states have case law that a parent's mental health is automatically at issue in a custody dispute.⁹⁴ While some states protect a parent's mental health records under privi-

⁹² Tippins & Wittmann, *supra* note 8.

⁹³ *The Uniform Marriage and Divorce Act of 1970*, 5 FAM. L.Q. 205 (1981).

⁹⁴ See *Owen v. Owen*, 563 N.E.2d 605, 608 (Ind. 1990) ("Mother placed her mental condition in issue when she petitioned for and was granted custody under the original order, and that condition remains in issue for the purposes of custody questions during the children's minority"); *Armstrong v. Heilker*, 850 N.Y.S.2d 673, 47 A.D.3d 1104 (N.Y. App. Div. 2008) (noting, "although not determinative, the mental health of a parent is necessarily relevant in every custody or visitation proceeding").

lege statutes, even these states allow discovery of mental health records when a parent's mental condition has become an issue at trial as part of an element of a claim or defense.⁹⁵ Many courts have found a parent's mental health to be the dispositive issue in determining custody.⁹⁶ Thus, it is usually imperative that evaluators provide information to the court not only about a parent's mental health, but more importantly, how the parent's mental health can and does impact their ability to care for the children and how it impacts their relationship with the child.⁹⁷

We disagree with Garber and Simon's ultimate conclusion that the best practice is to omit individual psychometric tests from custody evaluations. To the contrary, we believe psychological testing is one component of a CCE that can provide useful information in assisting evaluators, and ultimately, attorneys and the court, in determining what is in the best interest of the minor children. If an evaluator's opinion is solely based on interviews and record reviews, there is a greater risk of subjectivity. Psychological testing—using tests that have been normed on hundreds of thousands of subjects—provides a balance. Expert testimony based on personal opinion and cultural myths often has masqueraded as science.⁹⁸ Under *Daubert*, the testimony of the witness [is to be] 'more than subjective belief or unsupported speculation.'⁹⁹ The Wisconsin Court of Appeals wrote: "No matter how good experts' 'credentials' may be, they are 'not permitted to speculate.'¹⁰⁰ Psychological testing can provide information to support or disconfirm an evaluator's speculations.

⁹⁵ *Davidge v. Davidge*, 451 So.2d 1051 (Fla. Dist. Ct. App. 1984) (further distinguishing between a specific assertion of a mental condition in support of a claim from the general assertion of fitness typically made or implied in a child custody dispute).

⁹⁶ *See Maness v. Sawyer*, 950A.2d 830 (Md. Ct. Spec. App. 2008)(awarding sole legal and physical custody of children to the mother because they thrived in her care and the father's depression created "chaos" not only in the life of the mother but in the lives of the children).

⁹⁷ *Basham v. Wilkins*, 851 S.W.2d 491 (Ky. Ct. App. 1993)(past mental state and ability or inability to parent during that state is relevant to parent's ability to parent in the future because parent's mental ability may or may not have changed).

⁹⁸ GOULD, *supra* note 65, at 44.

⁹⁹ *Daubert*, 509 U.S. at 590.

¹⁰⁰ *State v. Giese*, 854 N.W.2d 687, 691 (Wis. 2014).

B. A Case Example

We conclude by providing a brief example of how psychological testing can assist an evaluator and the court in a child custody evaluation. In a complex custody dispute, which included allegations that the mother had a mood disorder and the father was physically and verbally aggressive to the mother, there was little data outside of the statements of the parents to support the veracity of their allegations. The father claimed the mother would start screaming during a dispute and would want to go outside and yell the information to neighbors. The father acknowledged he would restrain her from doing so, in order to prevent neighbors from knowing about the dispute. He did not believe he would scream and yell at her in the manner she described. Rather, he described them as normal marital arguments. The father further claimed the mother had mood related issues for years. The mother claimed the father would leave bruises on her when he would restrain her. While there were some photos she provided, the father claimed he does not know if they occurred during the times he restrained her and believed several had nothing to do with him. The mother further alleged the father would argue with police. However, there were no police reports documenting this.

In this instance, both parents completed a series of psychological tests, including the Minnesota Multiphasic Personality Inventory – 2 (MMPI-2), the Minnesota Multiphasic Personality Inventory – 2 – Restructured Format (MMPI-2-RF), the Personality Assessment Inventory (PAI), and Domestic Violence Inventory. Both parents completed them in an open and honest manner. The results showed no suggestion of any type of mood issue for the mother. Given that psychological testing can be used to generate hypotheses, as well as to assist in confirming/disconfirming hypotheses, the data from this mother's testing supported the hypothesis that there was not evidence of a mood related issue, although she acknowledged getting upset when her husband would become aggressive. On the other hand, the father, who also completed the tests in an open and honest manner, provided useful data. On the MMPI-2 he obtained a score of 71 on scale 4, the Psychopathic Deviate scale (PD). The Harris-Lingoes subscales for the PD scale shows he had a score of 67 on Authority Problems. On RC6 (Restructured Clinical scale 6),

Ideas of Persecution, he obtained a score of 65. Furthermore, on the Content scales, he had one elevation, which was on Antisocial Practices where he scored a 65. On the Supplementary scales, he also had one somewhat elevated score. He obtained a 65 on the Dominance scale. On the MMPI-2-RF he obtained an elevation on the Ideas of Persecution scale as well as the Juvenile Conduct Problems scale. Finally, on the PAI, he had an elevation on the Antisocial Features scale, where he obtained a score of 80, which equates to the 99.9th percentile. Furthermore, he scored an 82 on the Antisocial Behaviors subscale. An item analysis was conducted on all of the scales to specifically address what caused the scales to be elevated. On the DVI, the father scored in the moderate risk range for violence, suggesting that while he is not violence oriented, he can be impatient, easily annoyed and may be verbally aggressive. He also scored in the moderate risk range on a scale measuring his ability to cope with stress. This suggests that when he is under increased stress his coping abilities may become impaired. The end result is that in this matter, the pattern of scores for the father regarding antisocial behavior and attitudes was consistent with what the mother had reported. The testing, although not conclusive by itself, provided support for the mother not having any mood related issues as the father alleged and supported the mother's statements regarding the father's behavior toward her.

V. Conclusions

When a new idea arises to challenge existing practice, the proponent of the new idea should demonstrate how the new idea provides a similarly reliable and valid means of data gathering compared to the idea they wish to replace. Garber and Simon failed to demonstrate how the exclusion of data gathered through the multiple procedures reflected in the Forensic Model applied to child custody evaluations¹⁰¹ provides a more reliable and relevant set of information from which trustworthy and admissible expert testimony may be based. They provide no information demonstrating how their approach to custody evaluations and use of psychological results provide a more robust, trustworthy, and relevant set of information upon which expert opinions

¹⁰¹ Martindale & Gould, *supra* note 48.

may be based compared with the current standard of practice. Garber and Simon provide no empirical basis or legal foundation for their argument that exclusion of psychological test data improve the quality of information used by evaluations in developing hypotheses about parent behavior, co-parenting behavior, or how children may be affected by a parent's behavior, attitudes, and personality characteristics.

Expert testimony based upon ideas that have no basis in the scientific and professional knowledge of the discipline are, by definition, legally insufficient and therefore are inadmissible. Garber and Simon failed to address how the use of their "fit" model and the exclusion of psychological test data from the methodology used by child custody evaluators is a more reliable methodology than the current forensic model. They also fail to show how expert opinions based upon their untested methodology, that itself is based upon unscientifically unsupportable assumptions, would be admissible in court.

Garber and Simon provide no counterargument to the usefulness of psychological testing as described above. They fail to explain how their "fit" model provides data robust enough to replace the unique and comprehensive information psychological test data may provide. These arguments about the usefulness of psychological testing in child custody evaluation have been readily available for years in the peer-reviewed psychological and legal literature.¹⁰²

Garber and Simon's argument is based upon a myopic understanding of the complexity of child custody issues, assigning weight only to the dynamic interactions of parent-child behavior and the inferences of parent-child fit one can draw from direct parent-child observations and third party information. Their argument to do away with psychological testing and place the lion's share of an expert opinion about custodial fitness on direct observations and third party data is regressive. It places great weight on the accuracy of clinical observations of the evaluator and the accuracy of third party observers. The Garber and Simon ideas appear to place front and center great weigh on the *ipse dixit* of the expert witness and his or her observational acu-

¹⁰² Mary Johanna McCurley, Kathryn J. Murphy & Jonathan W. Gould, *Protecting Children from Incompetent Forensic Evaluations and Expert Testimony*, 19 FORENSIC EVALUATIONS 277 (2005).

men based upon their years of professional experience. We advise the reader to beware of the wozles in the Shoe Doesn't Fit and not to fall for them. The *Daubert* Court and its progeny at the federal court level and the forty plus state courts that have adopted the *Daubert* standard made clear the courts' view of expert opinions based upon nothing more than the expert's subjective determinations. Offering a model of child custody assessment that emphasizes unsupported expert testimony is out of step with the legal requirements of the work of child custody evaluators.