
FULL-LENGTH ARTICLE

Testing in Child Custody Evaluations—
Selection, Usage, and Daubert Admissibility:
A Survey of Psychologists

James N. Bow, PhD
Jonathan W. Gould, PhD
James R. Flens, PsyD
Dana Greenhut, PhD

James N. Bow received his PhD from The University of Michigan. He is Director of Psychology, Hawthorn Center and Adjunct Assistant Professor, Wayne State University, School of Medicine, Department of Psychiatry and Behavioral Neurosciences. His research interests involve child custody.

Jonathan W. Gould received his PhD from State University of New York, Albany. He has been in private practice since 1984 and is presently engaged in forensic practice and consultation as a member of Child Custody Consultants, a nationwide consulting consortium of specialists in child custody and other family law related matters. He has published widely in the child custody field, including his book, *Conducting Scientifically Crafted Child Custody Evaluations (2nd edition)*.

James R. Flens received his PsyD from Indiana State University. He is engaged in private practice in Brandon, FL that focuses on clinical and forensic psychology. He is also a member of the Child Custody Consultants, a nationwide consulting consortium of specialists in child custody and other family law related matters. His research interests include psychological testing in child custody evaluations.

Dana Greenhut received her PhD from University of Detroit-Mercy. She was a pre-doctoral psychology intern, Hawthorn Center when this study was conducted. She presently works for Psychiatry and Behavioral Medicine Professionals, Wayne State University.

Address correspondence to: James N. Bow, Hawthorn Center, 18471 Haggerty Road, Northville, MI 48168 (E-mail: bow@michigan.gov).

Journal of Forensic Psychology Practice, Vol. 6(2) 2006
Available online at <http://www.haworthpress.com/web/JFPP>
© 2006 by The Haworth Press, Inc. All rights reserved.
doi:10.1300/J158v06n02_02

ABSTRACT. A survey of psychologists (N = 89) was conducted to examine issues regarding test selection, usage rates, and Daubert admissibility in child custody evaluations. Findings revealed that respondents used selection criteria commonly cited in the forensic literature. They viewed the major purposes of testing as ruling out psychopathology and assessing personality functioning. Interestingly, less emphasis was placed on generating and testing hypotheses. In general, a limited number of tests and inventories were endorsed as meeting the Daubert standard. Implications for practice are addressed, particularly strategies for defending child custody testing practices against Daubert challenges. *[Article copies available for a fee from The Haworth Document Delivery Service: 1-800-HAWORTH. E-mail address: <docdelivery@haworthpress.com> Website: <<http://www.HaworthPress.com>> © 2006 by The Haworth Press, Inc. All rights reserved.]*

KEYWORDS. Daubert, child custody, psychological testing, family court, admissibility, forensic evaluations

Psychological tests are widely used in child custody work (Ackerman & Ackerman, 1997; Bow & Quinnell, 2001, 2002; Keilen & Bloom, 1986; and Quinnell & Bow, 2001). However, the admissibility of testimony regarding such tests is dependent on legal criteria. For years, the Frye standard (Frye v. United States, 1923) of “general acceptance” was used in both the Federal and State Courts as the standard for determining the admissibility of expert testimony. In 1993, the Daubert standard was introduced (Daubert v. Merrell Dow Pharmaceuticals, Inc., 1993), followed by other U.S. Supreme Court decisions such as General Electric Co. v. Joiner, 1997; and Kumho Tire Co. v. Carmichael, 1999. These legal decisions significantly impacted the threshold for admissibility of evidence in Federal courts, as well as the scrutiny placed on such evidence. Further, the adoption of Daubert or a Daubert-like standard by the majority of states clearly demonstrates its importance in the legal arena (Hamilton, 1998).

Prior to Daubert, psychological testing in forensic evaluations received little scrutiny and few challenges. The general acceptance standard was easier to attain. Psychologists often based their test selection on frequency of usage in the custody field, tests learned in graduate school and utilized in clinical practice, or tests commonly cited in journal articles or books. When a majority of child custody evaluators used

a test or technique, it was often viewed as accepted within the field regardless of its validity, reliability, or relevancy to the legal issue. Further, tests used for clinical purposes were commonly adopted for forensic practice, although the purpose, setting, type of decision-making, and level of test scrutiny vary in these settings (Otto & Heilbrun, 2002).

The introduction of the Daubert standard changed the landscape for forensic psychology. Daubert identified scientific knowledge as grounded in methods and procedures of science, with appropriate validation. Daubert outlined four guidelines for use by trial judges to assist them in determining the admissibility of expert testimony: (a) the underlying theory or technique has been tested or falsifiable, (b) the underlying theory or technique has been subjected to peer review and publication, (c) the underlying theory or technique has a known or potential error rate and maintenance of standards controlling the technique's operation, and (d) the underlying theory or technique is generally accepted (Daubert v. Merrell Dow Pharmaceuticals, Inc., 1993). Further, the testimony must be both scientifically valid and relevant. Overall, Daubert enhanced the role of the trial judge as the gatekeeper for expert testimony.

Daubert was further clarified by two ensuing Supreme Court decisions. In 1997, the Joiner decision (General Electric Company v. Joiner, 1999) held that the admissibility of expert testimony is reviewed under the "abuse of discretion" criteria. Consequently, the focus is on the legal reliability underlying the method used by the expert in reaching the opinion, rather than on the accuracy of the opinion itself. Also, this ruling further restated the role of the judge as the gatekeeper for expert testimony. The Kumho (Kumho Tire Co. v. Carmichael, 1999) decision expanded the focus of Daubert beyond the "scientific in nature" to include technical and other specialized knowledge. Consequently, Daubert was deemed relevant to psychological testimony; thereby eliminating any debate over this issue. The Kumho decision also focused attention on the word "knowledge" as the center of inquiry. Furthermore, it expanded the judge's gate-keeping function by allowing for broader latitude in decision making.

These Supreme Court decisions have been interpreted to require psychologists to go beyond the basic general acceptance standard and to focus more on scientific factors (e.g., test validity and reliability) and relevancy. These court decisions necessitate that psychologists be more careful and critical in their selection of child custody tests, techniques, and inventories. Heilbrun (1992), Marlowe (1995), and Otto, Edens, and Barcus (2001) have provided guidance in this regard. Psy-

chologists also have an ethical duty to keep abreast of the current professional literature and to present information to the courts utilizing reliable methods and procedures (Committee on Ethical Guidelines for Forensic Psychologists, 1991). Further, it is critical that psychologists conducting child custody evaluations not over-interpret or inappropriately interpret assessment data, and that they cautiously and conservatively address questions about reliability and validity (American Psychological Association, 1994).

Prior to Daubert, psychologists began to call upon the field to adopt more stringent criteria for choosing tests used in court evaluations. Heilbrun (1992) identified seven guidelines for forensic test selection: (1) commercially available with a manual, and listed/reviewed in *Mental Measurement Yearbook* or similar source, (2) standard guidelines for administration, (3) reliability coefficient exceeding 0.80, (4) relevancy to the legal issue or psychological construct underlying the legal issue, with available validation research, (5) application to the population and purpose for which the test was designed, (6) preference for objective tests and actuarial data combination, and (7) response style should be explicitly assessed. Otto et al. (2000) later adapted Heilbrun's (1992) criteria for choosing psychological tests in child custody evaluations.

As a result of the Daubert decision, Marlowe (1995) developed a hybrid model that blended scientific and legal principles. He also identified a set of psycho-legal criteria for use in determining the appropriateness of a test or instrument for forensic purposes. First, the content domain must sample all relevant areas with an ample range of item difficulty and moderate discrimination power. Second, the test or instrument must have a standardized administration procedure and justified norms. Third, the test or inventory must have reliability for the relevant population. Fourth, the test or inventory should be constructed in accordance with professional ethical standards and must have a manual describing validity, reliability, and normative data.

Lally (2003) surveyed forensic diplomates from the American Board of Forensic Psychologists (ABFP) about test usage and admissibility in six areas of forensic practice. However, he excluded child custody practice and used the Frye standard of general acceptance as the admissibility criteria. Although general acceptance is one of the Daubert criteria, Daubert delineated three other important criteria and made clear that other relevant criteria may be used by the judge. Further, Daubert or a Daubert-like standard has been accepted by the majority of states (Bernstein & Jackson, 2004), increasing its relevance in state courts.

At the present time, a variety of issues need to be explored regarding testing practice in child custody evaluations. First, what is the primary purpose of psychological testing and who should be tested? Second, what factors do psychologists consider in selecting tests/instruments? Third, how do psychologists view the admissibility of tests/instruments in child custody cases in accordance with the Daubert standard? Also, what impact, if any, does this have on test usage rates? The present study investigated these issues by surveying a sample of psychologists in the child custody field.

METHOD

The names of child custody evaluators were obtained from a number of sources, including an Internet search and referral lists, a child custody listserv, the public access site for the American Board of Forensic Psychology, and some evaluators known to the authors through conferences, publications, and professional activities.

An eight-page survey was developed to analyze important aspects of test selection and usage in child custody evaluations, particularly their admissibility according to Daubert. A cover letter, informed consent sheet, blank survey, and stamped return envelope were sent to all potential participants, who were asked to anonymously complete and return the survey. Potential participants were informed that all data would be coded and analyzed on a group basis to protect confidentiality. If they no longer conducted child custody evaluations or did not utilize testing in such evaluations, they were asked to return the blank survey indicating so. Those interested in the findings of the study were instructed to e-mail the first author requesting a summary of the results. Approximately one month after the initial mailing a reminder letter was sent to potential participants.

Of the 300 surveys mailed to potential participants, 135 were returned, making an adjusted return rate of 35%, which is comparable with past survey research. Of these, 89 of the respondents met the selection criteria (i.e., psychologists who conducted testing during child custody evaluation and completed the survey). Of the other returned surveys, 3 refused to complete the survey, 5 indicated they do not use testing in child custody evaluations, 9 no longer conducted child custody evaluations, 2 had retired from the field, 3 were non-psychologists, and 24 envelopes were undeliverable.

The vast majority of the respondents were male (70%). The average age was 54.13 (S.D. 6.53). Almost all were Caucasian (95.4%). Ninety-eight percent held a doctoral degree (80.9%–PhDs, 11.2%–PsyD, 5.6%–EdD, & 2.3%–MA). The vast majority was trained in clinical psychology (67.4%), followed by counseling psychology (18%), school psychology (4.5%), forensic psychology (2.2%), and other (7.9%). Twenty-one percent of the respondents had diplomates from the American Board of Professional Psychology, with forensic psychology as the most common specialty (71%). About one-quarter of forensic diplomates who received the survey completed and returned it. Respondents averaged 21.88 years in the mental health field, 17.81 years in forensic psychology, and 16.50 years in the child custody field. On average, they devoted 33% of their practice to child custody work, with a mean of 312 child custody evaluations in their career, and a mean of 14 evaluations over the past year. Almost all respondents worked in private practice (93.3%) in an urban/suburban area (91%). They represented 33 states, with the following distribution: 29.5% from the West, 29.5% from the Midwest, 17% from the East, and 24% from the South.

RESULTS

Respondents were queried about the individuals they tested in child custody evaluations. All respondents indicated that they tested the parents; whereas 69.3% tested adolescents and 51.1% tested children. Spouses and live-together partners were tested 64.8% of the time. In contrast, significant others and significant caretakers (e.g., grandparents, nannies, etc.) were only tested 19.3% and 10.2% of the time, respectively.

The primary purposes for administering psychological tests and inventories in child custody evaluations were explored. Among eight criteria offered, respondents were asked to select those that applied. Almost all respondents indicated that ruling out psychopathology was a primary purpose (96.6%), followed closely by assessing personality functioning (88.8%). A slight majority of respondents indicated that they used psychological tests and inventories to analyze parental strengths and weaknesses (62.9%) and test or generate hypotheses (57.3%). About one-third of respondents used psychological tests and inventories to confirm hypotheses (37.1%) and determine parenting capacity (32.6%). Further, findings indicated that the overwhelming majority of psychologists did not view psychological tests and inventories

as the primary data source (23.6%) or to determine the best interests of the child (24.7%).

In selecting tests and inventories, respondents were asked to rank-order twenty factors from 1 (most important) to 20 (least important), along with rating the importance of each factor on a Likert scale (1—not important to 7—extremely important). The findings are displayed in Table 1. The top five rankings were adequate validity research, sufficient body of research on the test/inventory, adequate reliability research, adequate normative sample, and acceptability in the child custody field. This was further reflected in the ratings, with 75% of the respondents rating these factors as 6s (highly important) or 7s (extremely important).

The next five factors focused on other critical areas often cited in the research as well. The bottom five rankings were computer scoring, cost, time to administer, availability of an interpretative report, and positive review in *Buros Mental Measurements Yearbook*. In addition, these factors received very low ratings with regard to importance.

Respondents were asked to rate a wide variety of tests, techniques, and inventories used in child custody evaluations according to the Daubert Standard. The following Daubert criteria were outlined in the survey: (a) the underlying theory can be tested, (b) the underlying theory has been subjected to peer review and publication, (c) the underlying theory has a known or potential error rate and maintains standards controlling the technique's operation, and (d) the underlying theory is generally accepted. The following rating system was used: (1) Recommended and Meets Daubert Standard, (2) Meets Daubert Standard, (3) Does Not Meet Daubert Standard, and (4) No opinion. The latter category was reserved for those unfamiliar with a test, inventory, or technique, or those who had no opinion about a test, inventory, or technique. Respondents were also asked to indicate if they typically used (> 90% of time) the test, inventory, or technique.

Table 2 shows those tests and inventories identified by the majority of respondents as meeting Daubert (i.e., 1—Recommended or 2—Not Recommended) and having a category 3 rating (Does Not Meet Daubert Standard) of less than 25%. The Minnesota Multiphasic Personality Inventory-2 (MMPI-2) received the highest Daubert rating (Meets Daubert-Recommended and Not Recommended = 95.2%), followed by the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III; 86.1%), Minnesota Multiphasic Personality Inventory-Adolescent Version (MMPI-A; 87.3%), Wechsler Intelligence Scale for Children-III/IV (86.1%), and Millon Clinical Multiaxial Inventory III (MCMI-III;

TABLE 1. Importance of Factors in Selecting Child Custody Tests and Inventories

Factor	Ave. Rating	S.D.
1. Adequate validity research	6.26	0.88
2. Sufficient body of research/publications	6.29	0.88
3. Adequate reliability research	6.19	1.03
4. Adequate normative sample	6.10	1.07
5. Accepted in CCE field	5.82	1.24
6. Relevance to the legal issue	5.98	1.15
7. Acceptability in scientific community	5.73	1.22
8. Standard administration procedures	6.14	0.95
9. Objective measure	5.73	1.22
10. Manual	5.65	1.47
11. Peer reviewed	5.63	1.14
12. Commercially available	5.17	1.80
13. CCE norms	4.80	1.42
14. Frequently used in survey data	5.00	1.50
15. Ease of administration	4.16	1.32
16. Computer scored	3.78	1.51
17. Cost	3.16	1.65
18. Time of administration	3.58	1.35
19. Interpretative report	3.29	1.63
20. Positive Buros review	3.37	1.60

Note. Factors are listed in rank order according to respondents' mean rankings on a 20-point scale (1 = most important and 20 = least important). Respondents also rated the factors on a Likert scale ranging from 1(not important) to 7 (extremely important).

77.6%). Of these tests, only the MMPI-2 (90.6%) and the MCMI-III (58%) were recommended and typically used by the majority of respondents.

The Personality Assessment Inventory (PAI), Psychopathy Checklist-Revised (PCL-R) and Beck Depression Inventory-2nd Edition (BDI-II) were endorsed as meeting the Daubert standard as well, but were infrequently used. The Rorschach Method-Comprehensive System (Exner, 1993) was the only projective personality measure seen as meeting the Daubert criteria, although only one-fourth recommended its use. Among parenting inventories, the Parenting Stress Index was

TABLE 2. Psychological Tests or Techniques Meeting the Daubert Standard

Specific Test of Technique	Percent of Respondents					% Typically Using
	Recommended	Not Recommended	Meets Daubert Standard	Does Not Meet Daubert	No Opinion	
<i>Intelligence Tests</i>						
Wechsler Adult Scale III	54.2	36.1	3.6	0	6.0	5.0
Wechsler Intelligence Scale for Children III/IV	46.8	39.2	0	0	13.9	11.0
Stanford-Binet Intelligence Scale—4th	34.2	30.3	2.6	2.6	32.9	1.3
<i>Academic Tests</i>						
Wide Range Achievement Test	24.0	52.0	4.0	4.0	20.0	9.1
Wechsler Individual Achievement Test	23.4	32.5	0	0	44.2	1.3
Woodcock-Johnson Battery	25.3	40.0	2.7	2.7	30.7	2.7
Peabody Individual Achievement Test	17.3	33.3	4.0	4.0	44.0	2.6
<i>Objective Personality Test or Technique</i>						
MMPI-2	77.4	17.9	1.2	1.2	3.6	90.6
MMPI-A	62.0	25.3	2.5	2.5	10.1	38.8
Millon Clinical Multiaxial Inventory	46.3	31.3	7.5	7.5	15.0	58.0
Millon Adolescent Clinical Inventory	34.7	28.0	5.3	5.3	32.0	19.7
Personality Assessment Inventory	32.9	22.8	2.5	2.5	41.8	17.9
Beck Depression Inventory II	18.2	37.7	14.3	14.3	29.9	11.5
Psychopathy Checklist-Revised	24.4	38.5	7.7	7.7	29.5	5.2
<i>Projective Personality Measures</i>						
Rorschach-Exner Scoring	25.6	33.3	24.4	24.4	16.7	--- ^a
<i>Parenting Inventories</i>						
Parenting Stress Index	24.4	37.2	10.3	10.3	28.2	27.0
<i>Children's Rating Scales</i>						
Child Behavior Checklist	29.3	29.3	8.3	8.3	33.3	23.0
Conners' Parent Rating Scale	19.7	30.3	9.2	9.2	40.8	32.0

Note. Each test or inventory was rated as: recommended and meets Daubert standard, meets Daubert standard, does not meet Daubert, or no opinion (i.e., unfamiliar with test/technique or have no opinion). Respondents were also asked to indicate if they typically ($\geq 90\%$ of time) use a test or technique.

^a Due to a printing error on the survey form, the usage response line for the Rorschach Comprehensive Scoring System was omitted.

the only instrument selected as meeting the criteria as well, with about one-fourth of respondents recommending it and typically using it in child custody evaluations.

Table 3 shows tests and techniques that almost half of the respondents rated as not meeting the Daubert standard. Projective drawing techniques, such as the Human Figure Drawing, House-Tree-Person Drawings, Kinetic Family Drawing were seen by the overwhelming majority of respondents (> 78%) as not meeting the Daubert standard, although they were typically used by roughly 10% of respondents. Classic apperception tests (e.g., Thematic Apperception Test and Children's Apperception Test) and Sentence Completion techniques were also seen as failing to meet the Daubert standard, although the latter was used by 25% of respondents. The Rorschach Method without the Comprehensive Scoring System was clearly viewed as failing to meet the Daubert standard, although when used with the Comprehensive Scoring System (see Table 2), it met such criteria.

None of the custody specific tests and techniques were viewed as meeting, or even approaching, the Daubert standard. These instruments combined Daubert percentages (e.g., Meets Daubert Standard-Recommended and Not Recommended) never exceeded 13, falling far below the 50% cut-off.

Table 4 lists the tests, inventories, and techniques that received a "no opinion" of approximately 50% or higher. Respondents were unfamiliar with these tests, inventories, or techniques, or had no opinion about them. This list involves numerous screening IQ tests and personality specific tests, along with some parenting inventories and rating scales.

DISCUSSION

Since the majority of states have adopted Daubert or a Daubert-like standard (Bernstein & Jackson, 2004), increased legal scrutiny will be applied to child custody evaluations by judges and attorneys practicing in those states that have adopted this admissibility standard. In particular, we believe there will be increased legal scrutiny about the use of psychological testing in child custody evaluations. Consequently, psychologists will need to be well versed in the properties of psychological testing in such evaluations, along with the legal issues that may arise from a Daubert challenge. The present study attempted to explore these issues through surveying psychologists about their test selection, usage, and views of Daubert admissibility. The respondents in this study (N =

TABLE 3. Psychological Tests and Measures Not Meeting Daubert Criteria

Specific Test or Measure	Percent of Respondents			No Opinion	% Typically Using
	Not Meeting Daubert Standard	Meets Daubert Standard Recommended	Meets Daubert Standard Not Recommended		
<i>Visual-Motor Measure</i>					
Bender Gestalt	48.0	5.3	18.7	28.0	3.8
<i>Personality Measures</i>					
Rorschach-Non-Exner Scoring	72.7	1.3	3.9	22.1	2.6
Thematic Apperception Test	77.2	1.3	7.6	13.9	5.2
Children's Apperception Test	74.4	1.3	6.4	17.9	9.1
Roberts Apperception Test	50.0	3.8	11.5	34.6	9.1
Human Figure Drawing	78.2	0	9.0	12.8	10.4
House-Tree-Person Drawings	78.9	1.30	5.3	14.5	10.4
Kinetic Family Drawing	82.1	0	6.4	11.5	17.9
Family Relations Test	45.5	0	7.8	46.8	7.9
Sentence Completion	71.1	1.3	13.2	14.5	25.3
<i>Custody Instruments</i>					
Bricklin Perceptual Scale	62.5	5.0	7.5	25.0	12.8
Perception of Relationship Test	64.1	5.1	7.7	23.1	11.5
Parent Awareness Skills Survey	68.4	1.3	7.9	22.4	7.8
Brinklin's ACCESS	57.7	2.6	3.8	35.9	1.3
ASPECT	49.4	1.3	10.4	39.0	3.9

Note. Each test or inventory was rated as: recommended and meets Daubert standard, meets Daubert standard, does not meet Daubert standard, or no opinion (i.e., unfamiliar with test/technique or have no opinion). Respondents were also asked to indicate if they typically ($\geq 90\%$ of time) use a test or technique

TABLE 4. Psychological Tests and Techniques with High No Opinion Rating

Specific Test or Technique	% No Opinion
<i>Intelligence Tests</i>	
Kaufman Brief Intelligence Test	51.9
Slosson Intelligence Test	63.6
Shipley Institute of Living Scale	62.3
Test of Nonverbal Intelligence	76.3
<i>Visual-Motor Measures</i>	
Development Test of Visual Motor Integration	56.9
<i>Personality Tests & Measures</i>	
NEO Personality Inventory-Revised	70.7
California Personality Inventory	64.5
Adolescent Psychopathology Scale	80.5
Reynolds Child/Adolescent Depression Scale	73.3
Zung Depression Scale	79.7
State-Trait Anger Expression Inventory	73.7
Spousal Assault Risk Assessment Guide	61.0
Trauma Symptom Inventory	52.6
Trauma Symptom Checklist for Children	61.3
Detailed Assessment of Posttraumatic Stress	82.4
<i>Substance Abuse Measures</i>	
Substance Abuse Subtle Screening Inventory	50.0
Michigan Alcohol Screening Test	55.8
<i>Parenting Inventories</i>	
Stress Index for Parenting Adolescents	74.7
Child Abuse Potential VI	46.1
Parent-Child Relationship Inventory	48.7
Parent Satisfaction Scale	80.8
Parent Alliance Measure	73.3
<i>Parent Rating Scales</i>	
Personality Inventory for Children	61.1
Child Sexual Behavior Inventory	67.1
Behavior Assessment System for Children	66.2
<i>Other Scales</i>	
Sexual Abuse Legitimacy Scale	73.0
Uniform Child Custody Evaluation System	51.9
Custody Quotient	51.3

Note. No opinion was given when a respondent was unfamiliar or had no opinion about a particular test or technique.

89) were highly trained and experienced psychologists in the child custody field, 21% of whom were diplomates of the American Board of Professional Psychology. However, the sample size of this study was relatively small, so caution must be used in generalizing the results.

Respondents viewed the major purpose of testing as a way of ruling out psychopathology and assessing personality functioning. Testing

was not viewed as the primary data source, or as a means of assessing parenting capacity or best interest standards. This approach is important because psychological data are one source of data collection, and should not be used in isolation (Gould, 1998, 2005; Heilbrun, 1992, 2001). Further, psychological tests are generally very limited in their ability to assess parenting capacity or specific best interest criteria. Another interesting finding was that only about half of the respondents identified psychological testing as a means of testing hypotheses, which is one of its major uses in forensic assessment (Gould, 1998, 2005; Heilbrun, 1992, 2001).

Data analysis was limited to psychologists who used psychological testing in child custody evaluations; this excluded a few psychologists who returned the survey stating that they did not conduct psychological testing as part of custody evaluations. All of the psychologists included in the study reported testing parents. The vast majority (69.3%) tested adolescents, but considerably fewer (51.1%) tested children. This was probably due to the greater availability of robust adolescent personality measures (e.g., MMPI-A, MACI). The use of objective testing with children is limited by requisite reading skills, usually 4th grade or higher, and the transparency of test questions. Of the personality tests rated by respondents as meeting the Daubert standard in this study, only the Rorschach Method with the Comprehensive Scoring System (Exner, 1993) was applicable to children. Spouses and live together partners were tested 64.8% of the time, which was higher than previous research reported (Bow & Quinnell, 2001). Nevertheless, it is surprising that this number is not much higher considering the integral roles these individuals play within the family.

In selecting tests and inventories, respondents gave the highest rankings and ratings to many of the factors cited in the research (Heilbrun, 2001; Marlowe, 1995; Otto et al., 2000), such as adequate validity and reliability, sufficient research/publication, adequate normative sample, acceptability in the scientific community, and relevance to the legal issue. These factors are also important for Daubert admissibility. Factors receiving low ratings and rankings involved less critical issues in a Daubert challenge, such as cost, ease of administration, computer scoring, time of administration, and availability of an interpretative report. Surprisingly, a positive review in the *Buros Mental Measurement Yearbook* was ranked and rated last. Ironically, *Buros* is viewed as an authoritative source in test review and its low standing in this study is perplexing.

Respondents selected relatively few tests and inventories as meeting the Daubert standard. The intelligence and academic tests chosen were well established and well known, but the typical usage rate was low. The latter supports prior research indicating that these tests are infrequently used in child custody evaluations (Quinnell & Bow, 2002; LaFortune & Carpenter, 1998).

Personality tests have greater utility in child custody evaluations, particularly with regard to ruling out psychopathology and assessing personality traits and characteristics, areas which respondents identified as most important. The highest percentage of respondents indicated that the MMPI-2 met the Daubert standard, with the vast majority recommending the test as well. Previous research (Ackerman & Ackerman, 1997; LaFortune & Carpenter, 1998; Quinnell & Bow, 2001) has affirmed its popularity in custody work. Also, child custody norms have been established by some researchers (Bathurst, Gottfried, & Gottfried, 1997; Flens, 2004; Strong, Greene, Hoopes, Johnston, & Olesen, 1999). However, defensive profiles on the MMPI-2 are common in child custody evaluations, which complicate its interpretation (Graham, 2000). Otto (2002) reported that court challenges to the admissibility of testimony involving the MMPI-2 are rare. Further, he noted that appellate court cases regarding the MMPI-2 found that none failed to meet the evidentiary standard set forth in Frye or Daubert when the focus was on emotional functioning or psychopathology.

The MCMI-III also received a high endorsement in this study for meeting the Daubert standard. However, much controversy has surrounded the use of the MCMI-III in forensic settings. Originally, the second version (MCMI-II) was recommended over the MCMI-III due to the latter's limited validation research (McCann & Dyer, 1996). Later, Dyer (1997) recommended the use of the MCMI-III in forensic settings after the revised manual indicated its suitability for forensic practice. However, Rogers, Salekin, and Sewell (1999, 2000) expressed concerns about the scientific validity and error rates of the instrument and argued against its use in such settings, questioning its admissibility according to the Daubert standard. Recent validation studies have supported the use of the MCMI-III (Dyer & McCann, 2000; Retzlaff, 2000; Schutte, 2001).

Concerns have also been expressed about the MCMI-III normative sample (i.e., clinical rather than "normal" sample). However, McCann and Dyer (1996) note that the MCMI-III normative group involved a significant number of "high conflict couples" receiving marital therapy, which makes it applicable for child custody cases. Recently, concerns

have been raised about the MCMI-III being gender biased in general, and potentially overpathologizing female custody litigants on specific scales in particular (Hynan, 2004; McCann, Flens, Campagna, Collman, Lazzaro, & Connor, 2001). Despite these controversies, the MCMI-III is being widely used in forensic settings, particularly in assessing personality traits and personality disorders. Child custody norms have been developed by McCann et al. (2001), and Flens (2004). The 58% usage rate in the current study supports findings from prior custody research (Bow, Quinnell, Zaroff, & Assemany, 2002; Quinnell & Bow, 2001).

Another adult objective test endorsed by the respondents as meeting the Daubert standard was the Personality Assessment Inventory (PAI), which has gained increased popularity over the last five years. Its appeal is its medium length (344 items), non-overlapping item scales, and low reading level (4th grade). In addition, it provides clinical and community norms. Flens (2004) has also compiled child custody norms.

For adolescents, the MMPI-A and MACI were both viewed by respondents as meeting the Daubert standard. The MMPI-A is the youth version of the MMPI-2 and was developed to assess a wide variety of psychopathology. However, the use of the MMPI-A is limited by its length (478 items) and fairly high reading level (6th grade+). It is also important to note that scant research is available on the admissibility of the MMPI-A in court. The MACI was developed to assess personality functioning and is an adolescent counterpart to the MCMI-III. It is relatively short (160 items), but a sixth grade reading level is required. Problems with the MACI involve excessive item overlap among scales and the dearth of empirical research on the instrument (McCann, 1999).

Two other objective personality inventories were endorsed as meeting the Daubert standard in child custody evaluations: Beck Depression Inventory-II (BDI-II) and Psychopathy Checklist-Revised (PCL-R). The BDI-II is a brief inventory that screens for depression. However, the transparency of the items and its screening nature limit its usefulness in forensic evaluations. The PCL-R is a widely used instrument and appears to meet Daubert admissibility standards (Gacono, Loving, Evans, & Jumes, 2002). Nevertheless, its usefulness in child custody evaluations is extremely limited because most examinees would not fit the PCL-R normative sample or Hare's (1998) definition of psychopathy. This is probably why only five percent indicated that they typically use the instrument in child custody evaluations. Therefore, although respondents endorsed the BDI-II and PCL-R as meeting the Daubert stan-

dard in child custody evaluations, their usefulness in such evaluations appears limited.

Among parenting inventories, only the Parenting Stress Index (PSI) was endorsed as meeting the Daubert standard. A major concern with the PSI is the normative sample (Yanez & Fremouw, 2004), which includes 2,633 mothers, but no fathers. Supplementary data were collected on 200 fathers, but were not included in the calculation of the norms. Although Yanez and Fremouw acknowledged that this problem rendered the first Daubert criterion (i.e., theory or technique can be tested) marginally adequate, they maintained that the instrument met the Daubert standard for overall admissibility. It is also interesting that the Parent-Child Relationship Inventory and Child Abuse Potential (CAP) Inventory failed to be endorsed by respondents as meeting the Daubert standard. In contrast, Yanez and Fremouw's (2004) review of the CAP found that it clearly met all Daubert criteria. This discrepancy appears to be a function of the high rate of "no opinion" responses given.

Among parent rating scales, only the CBCL and CPRS were endorsed as meeting the Daubert standard. However, a Daubert challenge has occurred over an expert's methodology in using the CBCL with a mentally retarded child in a sexual abuse case (*Gier v. Educational Serv. Unit 16*, 1995). Since the CBCL was not validated on a mentally retarded population and consists of few questions focusing on sexual abuse, the Court of Appeals supported the District Court's ruling that the plaintiff's expert had failed to conform to the Daubert standard. This ruling is critical because it demonstrates the importance of using tests and inventories with the appropriate normative population.

Among projective personality tests, only the Rorschach Inkblot Method using the Comprehensive Scoring System (e.g., Exner scoring) was endorsed as meeting the Daubert standard. Although some debate has surrounded the use of the Rorschach (Gacono, Evans, & Viglione, 2002; Grove & Barden, 1999; Grove, Barden, Garb, & Lilienfeld, 2002; McCann, 1998; Meyer, 2000; Ritzler, Erard, & Pettigrew, 2002; Wood, Lilienfeld, Garb, & Nezworski, 1999; Wood, Nezworski, Stejskal, & McKinzey, 2001), a survey of Rorschach workshop participants by Weiner, Exner, and Sciara (1996) indicated that Rorschach testimony was rarely challenged in court. In the 7,934 federal and state court cases that participants provided Rorschach testimony, only six cases (0.08%) resulted in admissibility challenges, with only one case (0.01%) resulting in the testimony being inadmissible.

Projective drawings (e.g., Human Figure Drawing, House-Tree-Person, and Family Drawing) and techniques (Thematic Apperception Test and Sentence Completion) received little endorsement in this study for meeting the Daubert standard. This is not surprising, considering that these projective drawings/techniques vary in administration, scoring, and interpretation, lack established validity and reliability, lack known error rates, and lack a scientific foundation and testability (Medoff, 2003). Lilienfeld, Wood, and Garb (2000) have expressed similar concerns and question their admissibility according to Daubert criteria. Despite all the criticisms leveled against these techniques, a small number of psychologists continue to use them in child custody evaluations.

Among custody instruments, such as the Bricklin scales and ASPECT (Ackerman and Schoendorf, 1992), none were viewed as meeting the Daubert standard. This is consistent with the strong criticism these instruments have received (Connell, 2005; Heinze & Grisso, 1996; Krauss & Sales, 2000; Otto et al., 2000).

A number of tests and instruments received a high percentage of no opinion responses. This means the respondents were unfamiliar with the test or inventory or had no opinion about it. Some of these tests were fairly well known, while others were unknown or rarely used. This response option was provided so respondents would not be forced to rate tests or instruments about which they lacked knowledge. Interestingly, many of these tests or inventories were of a screening nature or problem-specific, which probably limits their use in child custody evaluations.

In terms of the limitations of this study, the relatively small sample size suggests that caution should be used in generalizing the findings. Second, the endorsement of tests, techniques, and inventories as meeting or not meeting the Daubert standard might be more influenced by usage rates and popularity, rather than the test's actual psychometric properties and adherence to specific Daubert criteria. A recent study focusing on the use of the MMPI-2 and MCMI-II/III in child custody evaluations revealed serious concerns about psychologists' understanding of psychometric issues (Bow, Flens, Gould, & Greenhut, 2005). Third, since few Daubert challenges have actually occurred regarding forensic testing, psychologists may be inexperienced in addressing this issue, which impacts their knowledge and expertise in this area. Lastly, and most importantly, it is crucial that the findings from this study not be used as a point of reference for Daubert admissibility due to the above mentioned limitations.

Implications for Practice

The findings from the present study provide some preliminary guidance for psychologists in the selection and usage of psychological tests, techniques, and inventories for child custody evaluations. This will become increasingly important as Daubert challenges become more common on the state level. With the majority of states adopting Daubert or a Daubert-like standard instead of the Frye test of “general acceptance” (Hamilton, 1998), increased scrutiny will be applied to psychologists’ testimony in child custody evaluations. Consequently, psychologists will have to be better prepared to defend and provide a rationale for their test selection and usage.

In selecting tests and inventories psychologists need to be cognizant of important factors identified in this study and previously stressed in the literature (Heilbrun, 1992, 2001; Marlowe, 1995; Otto, 2000), such as a proven record of reliability and validity, sufficient body of research, adequate normative sample, and general acceptance in the child custody field. Further, other factors identified included relevance to the legal issue, acceptability within the scientific community, standard administration procedures, preference for objective tests, and presence of a manual. Failure to consider and utilize these factors in test selection could jeopardize their admissibility into evidence. It is further recommended that custody evaluators maintain a collection of literature regarding each test/instrument they use, including the most recent edition of the manual, current reviews in *Buros*, and articles that support and criticize the use of the test/instrument. It is more advantageous to learn about a particular test’s strengths and weaknesses in the safety of one’s office, rather than on the witness stand during brutal cross-examination by an adversarial attorney.

It is important that psychologists use testing in child custody evaluations to generate and test hypotheses, along with ruling out psychopathology and assessing personality traits. Failing to test/generate hypotheses creates potential for confirmatory biases and confirmatory distortion. Further, testability and falsifiability are rooted in Daubert criteria.

Psychologists should carefully consider the specific tests they administer in child custody evaluations. Tests and inventories must be selected for a specific purpose, and selection may vary from case to case depending on the specific issues and concerns raised. Further, when selecting tests or inventories, it is critical that each Daubert criteria be carefully considered. According to the present study, respondents

viewed relatively few tests and inventories as meeting the Daubert standard. Further, techniques such as projective drawings, sentence completion, and classic apperception tests, were overwhelmingly viewed as failing to meet the Daubert standard, which affirms criticism in the professional literature (Lilienfeld et al., 2000; Medoff, 2004). This was also true for custody tests, such as the Bricklin scales and ASPECT, which have been strongly criticized as well (Austin, in press; Connell, in press; Heinze & Grisso, 1996; Otto et al., 2000). Consequently, it is advised that these tests/techniques not be used in child custody evaluations.

It is hoped the information gleaned from this study will provide some guidance to psychologists and the court in better understanding test selection, usage, and Daubert admissibility issues. Through improving testing practice, the quality of child custody evaluation will improve, along with the services provided to families within the court system.

REFERENCES

- Ackerman, M. J., & Ackerman, M. (1997). Custody evaluation practice: A survey of experienced professionals (revisited). *Professional Psychology: Research and Practice, 28*, 137-145.
- Ackerman, M. J., & Schoendorf, K. (1992). *Ackerman-Schoendorf Scales for Parent Evaluation of Custody (ASPECT)*. Los Angeles: Western Psychological Services.
- American Psychological Association. (1994). Guidelines for child custody evaluations in divorce proceedings. *American Psychologist, 49*, 677-680.
- Bathurst, K., Gottfried, A., & Gottfried, A. (1997). Normative data for the MMPI-2 in child custody litigation. *Psychological Assessment, 9*, 205-211.
- Bernstein, D.E., & Jackson, J.D. (2004). The *Daubert* Trilogy in the States. *Jurimetrics Journal, 44*, <http://ssrn.com/abstract=498786>
- Bow, J. N., & Quinnell, F. A. (2001). Psychologists' current practices and procedures in child custody evaluations: Five years after the American Psychological Association guidelines. *Professional Psychology: Research and Practice, 32*, 261-268.
- Bow, J. N., & Quinnell, F. A. (2002). A critical review of child custody evaluation reports. *Family Court Review, 40*, 164-176.
- Bow, J. N., Quinnell, F. A., Zaroff, M., & Assemany, A. (2002). Assessment of sexual abuse allegations in child custody cases. *Professional Psychology: Research and Practice, 33*, 566-575.
- Bow, J. N., Flens, J., Gould, J. W., & Greenhut, D. (2005). *An analysis of administration, scoring, and interpretation of the MMPI-2 and MCMI-II/III in child custody evaluations*. Manuscript submitted for publication.
- Committee on Ethical Guidelines for Forensic Psychologists. (1991). Specialty guidelines for forensic psychologists. *Law and Human Behavior, 15*, 655-665.

- Connell, M. (2005). Review of "The Ackerman-Schoendorf Scales for Parent Evaluations of Custody." *Journal of Child Custody*, 2, 195-209.
- Daubert v. Merrell Dow Pharmaceuticals, Inc. 509 U.S. 579 (1993).
- Dyer, F. J. (1997). Application of the Millon inventories in forensic practice. In T. Millon (Ed.), *The Millon Inventories: Clinical and Personality Assessment* (pp. 124-139). New York: Guilford.
- Dyer, F. J. (2000). The Millon clinical inventories, research critical of their forensic application and Daubert criteria. *Law & Human Behavior*, 24, 487-497.
- Exner, J. E. (1993). *The Rorschach: A comprehensive system*. Vol. 1. New York: Wiley.
- Flens, J. R. (2004, October). *Advanced institute on psychological testing*. Pre-symposium institute conducted at the fourth annual Association of Family and Conciliation Courts Child Custody Symposium, Nashville, TN.
- Frye v. United States, 293F. 1013 (1923).
- Gacono, C. B., Evans, F. B., & Viglione, D. J. (2002). The Rorschach in forensic practice. *Journal of Forensic Psychology Practice*, 2, 33-53.
- Gacono, C. G., Loving, J. L., Evans, F. B., & Jumes, M. T. (2002). The Psychopathy Checklist Revised: PCL-R testimony and forensic practice. *Journal of Forensic Psychology Practice*, 2, 11-32.
- General Electric Co. v. Joiner, 522 U.S. 136 (1997).
- Gier v. Educational Serv. Unit 12, 66 F.3d 940 (1995).
- Gould, J. W. (1998). *Conducting Scientifically Crafted Child Custody Evaluations*. Thousand Oaks, CA: Sage.
- Gould, J. W. (2005). *Conducting Scientifically Crafted Child Custody Evaluations (2nd Edition)*. Sarasota, FL: Professional Resource Press.
- Graham, J. R. (2000). *MMPI-2: Assessing Personality and Psychopathology, (3rd Edition)*. New York: Oxford.
- Grove, W. M., & Barden, R. C. (1999). Protecting the integrity of the legal system. *Psychology, Public Policy, and Law*, 5, 224-242.
- Grove, W. M., Barden, R. C., Garb, H. N., & Lilienfeld, S. O. (2002). Failure of Rorschach-Comprehensive-System-based testimony to be admissible under the Daubert-Joiner-Kumho standard. *Psychology, Public Policy, and Law*, 8, 216-234.
- Hamilton, H. G. (1998). The movement from Frye to Daubert: Where do the states stand? *38 Jurimetric: The Journal of Law, Science, & Technology*, 201-213.
- Hare, R. (1998). Psychopaths and their nature: Implications for the mental health and criminal justice system. In T. Millon, E. Simonsen, M. Birket-Smith, & R. Davis (Eds.), *Psychopathy: Antisocial, Criminal, and Violent Behavior* (pp. 188-223). New York: Guilford Press.
- Heilbrun, K. (2001). *Principles of Forensic Mental Health Assessment*. New York: Kluwer Academic/Plenum Publishers.
- Heilbrun, K. (1992). The role of psychological testing in forensic assessment. *Law and Human Behavior*, 16, 257-272.
- Heinze, M. C., & Grisso, T. (1996). Review of instruments assessing parent competencies used in child custody evaluations. *Behavioral Sciences and the Law*, 14, 293-313.
- Hunsley, J., Lee, C. M., & Wood, J. M. (2003). Controversial and questionable assessment techniques. In S. O. Lilienfeld, S. J. Lynn, & J. M. Lohr (Eds.), *Science and Pseudoscience in Clinical Psychology*. New York: Guilford Press.

- Hynan, D. (2004). Unsupported gender differences on some personality disorder scales of the Millon Clinical Multiaxial Inventory–III. *Professional Psychology: Research and Practice, 35*, 105-110.
- Keilin, W. G., & Bloom, L. J. (1986). Child custody evaluation practice: A survey of mental health professionals. *Professional Psychology: Research and Practice, 17*, 338-346.
- Krauss, D. A., & Sales, B. (2000). Legal standards, expertise, and experts in the resolution of contested child custody cases. *Psychology, Public Policy, and Law, 6*, 843-879.
- Kumho Tire Company v. Carmichael, 526 U.S. 137 (1999).
- LaFortune, K. A., & Carpenter, B. N. (1998). Custody evaluations: A survey of mental health professionals. *Behavioral Sciences and the Law, 16*, 207-224.
- Lalley, S. J. (2003). What tests are acceptable for use in forensic evaluations? A survey of experts. *Professional Psychology: Research and Practice, 34*, 491-498.
- Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The scientific status of projective techniques. *Psychological Science in the Public Interest, 1*, 27-66.
- Marlowe, D. B. (1995). A hybrid decision framework for evaluating psychometric evidence. *Behavioral Sciences and the Law, 13*, 207-228.
- McCann, J. T. (1998). Defending the Rorschach in court: An analysis of admissibility using legal and professional standards. *Journal of Personality Assessment, 70*, 125-144.
- McCann, J. T. (1999). *Assessing adolescents with the MACI: Using the Millon Adolescent Clinical Inventory*. New York: Wiley.
- McCann, J. T. (2002). Guidelines for forensic application of the MCMI-III. *Journal of Forensic Psychology Practice, 2*, 55-69.
- McCann, J. T., & Dyer, F. J. (1996). *Forensic Assessment with the Millon Inventories*. New York: Guilford.
- McCann, J. T., Flens, J. R., Campagna, V., Collman, P., Lazzaro, T., and Connor, E. (2001). The MCMI-III in child custody evaluations: A normative study. *Journal of Forensic Psychology Practice, 1*, 27-44.
- Medoff, D. (2003). The scientific basis of psychological testing: Considerations following Daubert, Kumho, and Joiner. *Family Court Review, 41*, 199-213.
- Meyer, G. J. (2000). On the science of Rorschach research. *Journal of Personality Assessment, 75*, 46-81.
- Otto, R. K. (2002). Use of the MMPI-2 in forensic settings. *Journal of Forensic Psychology Practice, 2*, 71-91.
- Otto, R. K., Eden, J. F., & Barcus, E. H. (2000). The use of psychological tests in child custody evaluations. *Family and Conciliation Courts Review, 38*, 312-340.
- Otto, R. K., & Heilbrun, K. (2002). The practice of forensic psychology. A look toward the future in light of the past. *American Psychologist, 57*(1), 5-18.
- Quinnell, F. A., & Bow, J. N. (2001). Psychological tests used in child custody evaluations. *Behavioral Science and the Law, 19*, 491-501.
- Retzlaff, P. D. (2000). Comment on the validity of the MCMI-III. *Law & Human Behavior, 24*, 499-500.
- Ritzler, B., Erard, R., & Pettigrew, G. (2002). *Protecting the integrity of the Rorschach expert witness*. A reply to Grove and Barden (1999) Re: The admissibility of testimony under Daubert/Kumho analyses. *Psychology, Public Policy, and Law, 8*, 201-215.

- Rogers, R., Salekin, R. T., & Sewell, K. W. (1999). Validation of the Millon Clinical Multiaxial Inventory for Axis II disorders: Does it meet the Daubert standard? *Law and Human Behavior, 23*, 425-443.
- Rogers, R., Salekin, R. T., & Sewell, K. W. (2000). The MCMI-III and the Daubert standard: Separating rhetoric from reality. *Law and Human Behavior, 24*, 501-506.
- Schutte, J. W. (2001). Using the MCMI-III in forensic evaluations. *American Journal of Forensic Psychology, 19*, 5-20.
- Strong, D. R., Greene, R. L., Hoppe, C., Johnston, T., & Olesen, N. (1999). Taxometric analysis of impression management and self-deception on the MMPI-2 in child custody litigants. *Journal of Personality Assessment, 73*, 1-18.
- Weiner, I. B., Exner, J. E., & Sciara, A. (1996). Is the Rorschach welcomed in the courtroom? *Journal of Personality Assessment, 67*, 422-424.
- Wood, J. M., Nezworski, M. T., Stejskal, W. J., & McKinsey, R. K. (2001). Problems of the Comprehensive System for the Rorschach in forensic settings: Recent developments. *Journal of Forensic Psychology Practice, 1*, 89-103.
- Yanez, Y. T., & Fremouw, W. (2004). The application of the Daubert standard to parental capacity measures. *American Journal of Forensic Psychology, 22*, 5-28.

RECEIVED: 01/13/05

REVISED: 04/07/05

ACCEPTED: 05/19/05