

parent, (5) each parent's parenting skills, (6) each parent's openness towards the child's contact with the other parent, (7) the parents' pre-separation caretaking and parenting roles, (8) the parents' expressed anger and bitterness regarding the divorce, (9) the parents' sexual orientation, and (10) the stated preferences of a 5-year-old child.

Ten years later, Ackerman and Ackerman (1997) surveyed 800 doctoral-level psychologists who conducted child custody evaluations and obtained usable responses from 201 (25%). Respondents spent 21 hours per evaluation—similar to the earlier survey—but these respondents reported devoting more time to reviewing collateral materials and report writing. Intelligence tests and projective measures continued to be the instruments most frequently employed with children, and the MMPI/MMPI-2 remained the most frequently used assessment instrument for parents, followed by the Rorschach Inkblot Technique.

Many custody evaluators also reported using assessment instruments with children that were developed specifically for use in custody contexts (Ackerman & Ackerman, 1997). Over one third used the Bricklin Perceptual Scales (Bricklin, 1990a) while 16% used the Perception of Relationships Test (Bricklin, 1989). Fewer respondents (11%) used the Ackerman-Schoendorf Scales for Parent Evaluation of Custody (Ackerman & Schoendorf, 1992), the one custody-assessment measure designed for entire families and adults. Fewer than 10% used other custody-assessment measures, specifically, the Parent Awareness of Skills Survey (Bricklin, 1990b) and the Custody Quotient (Gordon & Peek, 1989). Other investigators (e.g., Bow & Quinnell, 2001; Gourley & Stolberg, 2000) have reported findings regarding test usage by custody evaluators similar to those detailed by Keilin and Bloom (1986) and Ackerman and Ackerman.

Like Keilin and Bloom (1986) before them, Ackerman and Ackerman (1997) also asked custody evaluators to rate the importance of various factors to issues of child custody. According to the custody evaluators, the ten most important, in descending order of significance, were (1) the substance abuse status of each parent, (2) the parents' parenting skills, (3) parental attempts at alienation, (4) the nature and quality of the child's emotional relationship with each parent, (5) the emotional or psychological stability of each parent, (6) each parent's openness toward the child's contact with the other parent, (7) the parents' history of compliance with the court during the separation, (8) the parents' pre-separation caretaking and parenting roles, (9) the stated preferences of a 15-year-old or older child, and (10) the parents' expressed anger and bitterness regarding the divorce.

THE (LIMITED) SCIENCE OF CUSTODY EVALUATIONS

State statutes regarding children's best interests help us understand at least some of the practices of custody evaluators. We could (and later do) question, for example, whether (or when) a parent's mental health or the wishes of a child should be a

central focus in child custody cases. Still, evaluators who assess such factors are following explicit legal guidelines. More difficult to explain and more problematic, however, are other aspects of evaluation practices including the widespread use of well-established measures with no clear relevance to the custody context (e.g., measures of intelligence), attempts to measure constructs created to apply to child custody decision making (e.g., "parent alienation syndrome"), efforts to identify "parent of choice" (e.g., the Bricklin Perceptual Scales), and the use of measures that a significant number of psychologists view with skepticism (e.g., the Rorschach Inkblot Technique).

We are dubious about many child custody evaluation practices, because of the absence of solid psychological science and of clear criteria to be predicted by psychological science. We also hold two much more fundamental questions about child custody evaluations: Why has society and the law placed such importance on a prediction about *psychological* factors in determining custody? And if the goal is to minimize children's psychological risk, might there be better roles for psychologists to play—both as practitioners and as scientists—in custody disputes? For now, however, we focus on the lack of scientific evidence to support many of the instruments and practices of mental health professionals who serve as custody evaluators.

Heilbrun, Rogers, and Otto (2002) described a three-category typology of assessment techniques used in forensic contexts, including custody evaluations. *Clinical assessment instruments* are those developed to assess psychological constructs, typically for intervention purposes (e.g., measures of intelligence, psychopathology, academic achievement). *Forensically relevant instruments* assess constructs that are psychological in nature but may be of particular relevance in forensic contexts (e.g., measures of response style, risk for criminal offending). Finally, *forensic assessment instruments* are specifically designed to assess psycho-legal constructs. Here we review evidence in regard to the third and first categories of assessment techniques. We do not consider forensically relevant instruments because none have been used widely by custody evaluators, although that may change (Posthuma, 2003). We also raise concerns about "parent alienation syndrome" and other constructs that have been created for, and asserted to have scientific standing in, the context of custody evaluations.

Forensic Assessment Instruments: No Scientific Support

In the past 15 years, psychologists have developed a number of forensic assessment instruments purporting to assess children's best interests in custody disputes (see Grisso, 2003). Our bottom-line evaluation of these measures is a harsh one: These measures assess ill-defined constructs, and they do so poorly, leaving no scientific justification for their use in child custody evaluations.

The most widely used forensic assessment instrument (Ackerman & Ackerman, 1997) is the Bricklin Perceptual Scales

(BPS), described as a projective measure of parents' competence, supportiveness, follow-up consistency, and possession of admirable traits (Bricklin, 1990a). Using a stylus and rating card, children rate each parent on 32 different activities considered to be relevant to these four capacities. The parent who receives the greater number of positive ratings is identified as the "Parent of Choice." Bricklin asserts that the nonverbal nature of the task (using a stylus rather than a verbal response) allows for the assessment of the child's "unconscious preferences," which are less likely to be subject to distortion due to social desirability or parental persuasion. However, the BPS has been criticized on numerous grounds: There is no support for claims that it assesses children's unconscious preferences or that responses are not subject to external influence; the developer permits variation from standard test administration; the measure samples a relatively narrow range of parenting domains; the developer has not provided basic norms and psychometric properties of the measure; and data regarding concurrent and predictive validity are either absent or unconvincing (Heinze & Grisso, 1996; Melton, 1995; Melton, Petrila, Poythress, & Slobogin, 1997; Otto & Edens, 2003; Otto, Edens, & Barcus, 2000; Shaffer, 1992).

Another measure used fairly frequently is the Perception of Relationships Test (PORT; Bricklin, 1989), a projective drawing that is described as measuring the "whole organism or gut-level responses a child has toward a parent [that] are much more reflective of what the child's actual interactions or experiences with that parent have been" (Bricklin, 1993, p. 1). Seven drawing tasks completed by the child are scored to identify the "Primary Caretaking Parent." Like the BPS, the PORT has been widely criticized. Objections include the incomplete and confusing manual, unclear administration and scoring guidelines, minimal reliability data, missing norms, and lack of validity data (Carlson, 1995; Conger, 1995; Heinze & Grisso, 1996; Melton et al., 1997; Otto & Edens, 2003; Otto et al., 2000).

Bricklin (1990b) describes another measure, the Parent Awareness Skills Survey (PASS), as a "clinical tool designed to illuminate the strengths and weaknesses in awareness skills a parent accesses in reaction to typical child care situations" (p. 4). The PASS consists of 18 childcare scenarios selected to represent caretaking of children of various ages. The parent's responses are followed up with questioning by the examiner as needed, and scoring is based on guidelines in the test manual. The PASS also has been criticized for basic shortcomings: the absence of norms, reliability and validity data, and clear scoring guidelines (Otto & Edens, 2003; Otto et al., 2000). Of particular concern is the developer's suggestion that "the evaluator, by virtue of appropriate training in psychology and/or child development, can apply his or her own standards in assigning the suggested scores. The PASS allows for wide latitude in scoring since its main purpose is to discover the relative (rather than absolute) strengths and weaknesses any individual or compared set of respondents manifest" (Bricklin, 1990b, p. 11).

The Parent Perception of Child Profile (PPCP; Bricklin & Elliott, 1991) is described as a measure of parents' understanding of a child's development and needs across eight areas: interpersonal relations, daily routine, health history, developmental history, school history, fears, personal hygiene, and communication style. Because parents who more accurately assess their child are assumed to be better parents, the PPCP requires the examiner to assess the accuracy of each parent's report, using vaguely defined criteria that include the examiner's and third-party informants' opinions. According to the manual, data need not be gathered in all eight categories, and the examiner can decide which issues are most critical for a particular child and parent. The PPCP has been criticized for its incomplete manual, lack of scoring directions, and absence of reliability and validity data (Otto et al., 2000; Otto & Edens, 2003).

Another instrument used by evaluators with some frequency, the Ackerman-Schoendorf Scales for Parent Evaluation of Custody (ASPECT), is purported to be "a clinical tool designed to aid mental health professionals in making child custody recommendations" (Ackerman & Schoendorf, 1992, p. 1). The ASPECT is not a test, but an assessment approach that aggregates data from the parent (an open-ended "Parenting Questionnaire," the MMPI-2, the Rorschach Inkblot Technique, and an intelligence test) and from the child (the Rorschach Inkblot Technique, an intelligence test, an academic achievement test, and a projective story). Measures were selected based on the developers' review of the literature, and test scores are used to calculate a "Parental Custody Index" (PCI) for each parent. The PCI is considered to indicate parenting effectiveness, and judgments about the parents are based on their relative PCI values. With rare exceptions (e.g., Brodzinsky, 1993), reviews of the ASPECT have been uniformly negative. Criticisms include the absence of a clear relationship between many of the measures and behavior relevant to custody; the failure to assess factors clearly deemed relevant to custody decisions; and an absence of important data regarding basic psychometric properties, including predictive validity (Arditti, 1995; Heinze & Grisso, 1996; Melton, 1995; Melton et al., 1997; Otto & Edens, 2003; Otto et al., 2000; Wellman, 1994).

In summary, all measures that purport to assess constructs directly relevant to child custody determinations suffer from significant limitations. In fact, no study examining the properties of these measures has ever been published in a peer-reviewed journal—an essential criterion for science and, in theory, for the courts. In our view, the absence of scientific support should preclude the use of any of these forensic assessment instruments for any purpose other than research. We even have doubts about the value of research using these measures, because it is hard to conceive of any psychological test that could measure all the factors that might be relevant to child custody (Shuman, 2002) or that might assess the best custody arrangements for children when the criteria for fulfilling children's best interests are so poorly defined (Emery, 1999b).

Clinical Assessment Instruments: Some Cautions in the Custody Context

Heilbrun et al. (2002) describe measures of intelligence, personality, psychopathology, and academic achievement as clinical assessment instruments. In contrast to forensic assessment instruments, we believe use of many of these measures is warranted in forensic assessment contexts to the degree that they offer reliable and valid assessments of relevant constructs identified in the law. We do, however, wonder about the routine use of measures such as IQ tests, which can add to the time and expense of a custody evaluation without holding a clear relevance to the issue before the court.

A greater concern is the validity of clinical assessment instruments in the custody context, as a number of considerations suggest the need for caution. For one thing, as in other forensic contexts, examinees may be less than candid in their responses, including on psychological tests. Tests that do not include measures of response style are particularly vulnerable to dissimulation, while tests with embedded measures of response style are not necessarily impervious to false reporting.

Whether the constructs assessed by the instrument are, broadly conceived, “states” or “traits” is another important issue. Assessments of characteristics that commonly change over time (e.g., parental depression) provide a weak basis for an evaluator to make claims about how a parent functioned in the past or will function in the future. Because families are evaluated during a period of high stress, moreover, evaluators also must be cautious about drawing inferences about functioning at some later, hopefully less stressful, point in time. Given the very nature of custody disputes and the context in which most custody evaluations occur, it is particularly important that the evaluator not assume that instruments assessing more enduring styles will not change in response to situational factors. The Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999) direct that “a test taker’s score should not be accepted as a reflection of lack of ability with respect to the characteristic being tested for without consideration of alternate explanations for the test taker’s inability to perform on that test at that time” (p. 43). The upheaval of divorce constitutes a reasonable “alternative explanation” that should certainly be considered when interpreting a test score.

We do not want to throw out the baby with the bathwater. There may be a role for clinical assessment instruments in some custody evaluation contexts. More specifically, to the degree that there is a psychological construct that is relevant to the issues at the heart of a custody matter and there are valid psychological measures of that construct available, use of such measures can be of some value. Examples of relevant things that may need to be determined in a custody case might include whether a child has a learning disorder that needs special attention, whether a mother suffers from depression that affects her ability to meet

her children’s emotional needs, or whether a father has a substance-abuse disorder that results in him placing the children in at-risk situations when in his care.

Projective Measures

Our concerns about clinical assessment instruments apply to highly structured, well-validated, and well accepted measures of intelligence, academic achievement, and psychopathology. These issues present the greatest concerns, however, for unstructured, projective measures, given questions that have been raised about even basic psychometric properties of such tests, including their reliability and validity. There is a considerable difference of opinion and ongoing, active debate regarding the general utility of projective measures such as the Rorschach Inkblot Technique (compare Wood, Nezworski, & Stejskal, 1997; Wood, Nezworski, Lilienfeld, & Garb, 2003 and Weiner, 1996; Meyer, 1997, 2001), Draw a Person, and Human Figure Drawings. The very existence of this debate, in combination with some of the specific criticisms and potential dangers in the custody context, lead us to suggest that such measures not be used in child custody evaluation contexts, or any other evaluation contexts for that matter.

We do not have the space, expertise, or the inclination to review the broad and polarized literature on projective tests in this monograph. Thus we only point to the extensive and serious controversy, and note this: Questions about the value of projectives or any other assessment technique need to be debated and answered by psychological scientists outside of the courtroom. It is naive to expect judges to make informed judgments about the psychometric adequacy of projective measures in the context of a custody hearing. We also are concerned about the potential for evaluators to assert that projective measures have scientific authority while the underlying empirical, legal, and values questions remain unanswered, precisely because the “test” is mysterious to lay observers and therefore potentially misleading or difficult to challenge. A nonexpert might feel competent challenging the relevance or the validity of a relatively straightforward measure like an IQ test or an MMPI-2. Yet, despite more significant concerns about its psychometric properties, results of a Rorschach may be more difficult to challenge precisely because of its more obscure source of material and scoring (Shuman, 2002).

Clinical Interviews

The clinical interview is another assessment technique that requires considerable caution when used as a measurement technique in custody evaluations. Interviewers may yield inferences that are reliable or unreliable, valid or invalid, but there are no structured interviews with well-established psychometric properties specifically developed for use in the child custody context, and survey data regarding psychologists’ custody evaluation practices indicate that use of any structured interview approach is virtually unheard of (Ackerman & Ackerman, 1997; Keilin & Bloom, 1986). Thus, differences between

interviewers may result from variance in the family's responses or from the contrasting structure, content, or interpretation of the interview. We urge psychological scientists to work to develop structured interviews for the custody context. In the meantime, we expect custody evaluators to continue to interview families. Although we are dubious about the psychometrics of unstructured interviews, we find some comfort in the fact that, unlike projective measures, interviews are more straightforward and understandable and hopefully are not presented as providing data as scientific-sounding as that of a test.

Direct Observation

Direct observation of parent-child interactions is another complex and generally unstandardized assessment strategy. Threats to validity include reactivity, unreliable coding systems, unrepresentative samples of behavior, and problematic data compilation and analysis. As with interviews, we urge the development of standardized observation measures for use in the custody context, and urge evaluators to describe their observations clearly and to identify the inferences they draw from observational assessments.

Combining Assessment Results and Drawing Inferences

This last point raises a broader and very important issue. All clinical assessment instruments assess constructs that, at most, are only indirectly relevant to custody; thus their use in custody evaluations typically requires inferences to be made. Once a parent's depression or a child's academic abilities are assessed, for example, the examiner may draw some inference regarding how that factor is relevant to the best interests of a child. The question is: How is the examiner to draw conclusions from a single measure or, even more importantly, combine data from several sources to form a conclusion about the best interests of the child? For example, how does an evaluator weigh the results of a hypothetically accurate (impossible in practice) evaluation where all data indicate that the mother is an effective disciplinarian but not terribly warm and that the father is warmly supportive but not good at setting limits? Thus, our concerns with clinical assessment instruments are not only how to measure relevant constructs reliably and validly in a difficult context, but also how to synthesize multiple measurements in a manner relevant to the ultimate issue of a custody determination. Perhaps ideally, the law would provide a formula for making such decisions, but the factors to be considered in the law are rarely even ranked relative to one another.

Controversial Topics Requiring Further Investigation

Surveys of practicing custody evaluators indicate that, in addition to using clinical assessment instruments and dubious forensic assessment instruments, they also frequently assess certain quite controversial constructs. We illustrate our concerns by focusing on three in particular: (a) parental alienation

syndrome (PAS), (b) children's wishes regarding custody, and (c) overnight visitation for very young children.

PAS: Asserting Science Where There Is None

"Parental alienation" is a construct ranked high on the list of factors evaluators consider to be directly relevant to custody decision making. There is no test instrument designed to measure parent alienation. Rather, it is a "diagnosis" reached through clinical interviews. Some experts have testified to making the diagnosis of parental alienation syndrome, and their testimony is claimed to be an important influence on judicial decision making (Gardner, 2004).

"Parental Alienation Syndrome" is a term created by psychiatrist Richard Gardner (2001) based on his clinical experience with custody disputes. Gardner asserts that PAS, which he says develops almost exclusively in the context of custody disputes, is characterized by one parent "programming" a child against the other parent (Gardner, 2001). The assumption is that a child's disdain for one parent is generally unjustified and solely attributable to denigration on the part of the other, alienating parent. Gardner (2004) also claims that PAS can be "diagnosed" reliably and validly by expert evaluators, although he offers no explicit criteria for doing so or objective evidence to support his claim (Emery, 2005).

We recognize that parents often undermine each other's relationships with their children following separation (Emery, 2005; Kelly & Johnston, 2001). We also note that many state statutes include a "friendly parent" rule, a preference for awarding custody to the parent who will be more likely to promote the children's relationship with the other parent (Elrod & Spector, 2004). However, the scientific status of PAS is, to be blunt, nil. As Gardner (2004) himself noted in a recent posthumous publication, only one study of parent alienation ever attempted a statistical analysis: his own. Very recently, Johnston conducted two studies of case records designed to identify the sources of alienation; she found many contributing factors leading to a child aligning with one parent against the other, including high-conflict custody litigation and poor parenting on the part of the "alienated" parent (cited in Johnston & Kelly, 2004).

We believe that it is blatantly misleading to call parental alienation a scientifically based "syndrome" (Emery, 2005). Careful assessments of each parent's willingness to support the other coparent clearly may be relevant to custody, but there is no established way of measuring "alienation." Evaluators therefore must carefully identify the sources of their information concerning a more or less "friendly" parent, as well as the inferences they draw from these assessments. Certainly, these assessments are best conducted by an evaluator who interviews both parents, something Gardner (2001) did not do in many cases.

Children's Wishes

Surveys indicate that custody evaluators place considerable importance on children's stated preferences regarding custo-

dy—particularly the preferences of adolescents, but also of children as young as 5 years old (Ackerman & Ackerman, 1997; Keilin & Bloom, 1986). This surely reflects the fact that children's wishes regarding custody typically are included in state laws as a factor to be considered when determining children's best interests. In fact, some statutes explicitly direct that the wishes expressed by a child of a given age—for example, 12 years old—should determine custody if there is no reason why those wishes should not be followed (Elrod & Spector, 2004). Although all agree that the wishes of teenagers can be influenced by unfortunate circumstances (e.g., a parent's greater material resources or permissiveness), laws regarding the expressed wishes of children of a certain age both respect the increasing autonomy of adolescents and recognize the realistic difficulty of trying to keep children in an arrangement to which both they and one parent object.

A policy of acting on the freely expressed wishes of an adolescent is not without problems, but far bigger problems (and controversies) arise in regard to wishes of children who (a) are school aged or even younger and/or (b) do not come forward with a freely expressed preference. Some psychologists have offered that, even in these circumstances, children should be encouraged to express a preference regarding custody as a means of empowering them (see Weithorn, 1987). Others express concern that, instead of giving children the right to have input, such policies give children the responsibility for making adult decisions—decisions that the adults have failed to make themselves (Emery, 2003). Still others say that children's preferences should be assessed only sensitively and indirectly and that this information should be used as feedback to facilitate independent parental decision making (McIntosh, Long, & Moloney, 2004).

One of us has taken a strong position against attempting to assess children's unexpressed wishes (Emery, 2003), but our present concern is more basic. The freely offered preferences of children—particularly older children—are important considerations in custody evaluations for both practical and legal reasons, but there is no direct evidence on how or indeed whether evaluators should assess the wishes of children who, for whatever reason, do not express them.

Overnights With Infants and Toddlers

A final controversy we will discuss is whether or to what extent infants and toddlers should have overnight visits with their nonresidential parents. Children's age in relation to overnights is not a consideration mentioned often in surveys of custody evaluators, but it stands as an example of the sorts of controversial issues that evaluators often are asked to address. Other such issues include the question of whether a residential parent with primary physical custody should be allowed (if there are good reasons) to move with the child away from a nonresidential parent, or under what circumstances parental conflict is so intense that joint physical custody is unworkable.

Using differing interpretations of attachment theory, leading psychological scientists have taken strong and very different positions on the issue of overnights involving young children. A document prepared for the Spokane (Washington) Bar Association, and endorsed by many leading attachment researchers, called attention to the psychological importance of young children's secure attachment with a primary attachment figure. Based on research and theory on the primary attachment, the report recommended against overnight visits with the nonresidential parent until children are 4 years old (Spokane County Bar Association, 1996). In contrast, in a paper published in a major family-court journal, other leading psychological scientists highlighted the importance of children's attachments to multiple caregivers. Focusing on the value of developing multiple attachments, the authors recommended that infants should have regular overnight visits with nonresidential parents in the first year of life (Kelly & Lamb, 2000). Both interpretations offered various caveats about the quality of children's relationship with the nonresidential parent, parental cooperation, and similar issues, but they clearly came to very different substantive conclusions about what psychological science indicates regarding whether, when, or how often infants and toddlers should have overnight visits with nonresidential parents.

There is only meager direct evidence on the harm (Solomon & George, 1999) or absence of harm (Pruett, Williams, Insabella, & Little, 2003) associated with overnight visits for very young children. As with the issue of children's wishes, the psychological scientists debating the question of overnight visits apparently come to logical conclusions based on their own, theoretical premises, yet the limited state of knowledge allows reasonable scientists to come to opposing conclusions. Such differences of opinion are of great value in science, but when translated into policy recommendations, they can confuse and confound judges, lawyers, evaluators, and parents. For example, we have had distraught mothers approach us in shock after being court-ordered to stop breast-feeding their infants to allow for smoother overnight visits, yet we also know of judges who claim to overturn consensual parenting plans if they include overnight visits for children 3 years of age or younger because of worries about disrupting attachments.

One of us has developed a set of guidelines for parents about overnights and other arrangements for young children that represents what we believe to be a balanced position (Emery, 2004). However, our point here is that, whatever conclusion one reaches, it is based on limited evidence. Psychological scientists need to recognize and acknowledge their limited data base.

Our bigger point, to which we turn shortly, is this: Custody decision making and custody evaluations have an impossible task in attempting to determine children's future "best interests" in cases where parents cannot agree. Neither the wisest judge nor the most insightful evaluator has good answers to impossible questions.

The custody report completed by Dr. Hagan in our fictional case illustrates our various concerns with the limited science of

custody evaluations and, more importantly, shows the problems that can arise under the regime of vague custody laws and vague professional and ethical standards for custody evaluators.

The Deer-Doe Case: Dr. Hagan's Custody Report

Dr. Hagan wrote a 35-page report summarizing his evaluation of the Deer-Doe family. The report contained precise details of the results of the various standardized tests, but the lawyers were only really interested in the final paragraphs under the heading, *Summary and Recommendations*.

"In summary, substantial evidence points to Ms. Deer-Doe's longstanding depression, her intense, repressed hostility toward Mr. Deer-Doe, and her alienation of the children against their father. In contrast, Mr. Deer-Doe appears to be well adjusted, is eager to promote the children's relationship with their mother, and is able and interested in being a full-time father. It therefore is recommended that, in order to promote his best interests, Carlos Deer-Doe be shifted immediately to his father's custody with regular visits with his mother, provided that she enters into individual psychotherapy.

"Although Isabella's intense anger at her father is largely a product of alienation, no change in custody is recommended for her at this point in time, because she is closely allied with her mother and is likely to continue to reject and rebel against her father's care. Instead, individual psychotherapy and family therapy with her father is recommended for Isabella, with further evaluation in 3 to 6 months depending upon the recommendations of Isabella's therapists and her mother's therapist, if relevant. A key consideration at that time will be whether Isabella's stated wish to live with her mother, if she continues to voice this preference, is a result of alienation."

When he read the evaluation, John Deer-Doe was jubilant. He felt vindicated, eager to be a full-time father again, and excited about the prospect of starting his new family. He vowed he was now going to get remarried "the day after my divorce is final." His lawyer, who also was encouraged by Dr. Hagan's report and recommendations, told John that the evaluation was not only a victory for him but for all fathers. "Sometimes the system really does work," she offered.

Jane Deer-Doe's reactions were understandably quite different. Shocked and panicked, she became emotionally distraught in her lawyer's office. He eventually helped Jane calm down by telling her that he had learned only recently that Dr. Hagan, who used to be fair and evenhanded, had become notoriously biased in favor of fathers as a result of losing custody in his own, bitter divorce. If he had known this a few months ago, Jane's lawyer told her, he never would have agreed to Dr. Hagan as the court-appointed evaluator.

Ms. Deer-Doe's attorney went on to offer that he would postpone the pending hearing in order to get a second evaluation by another mental health professional and have Dr. Hagan's evaluation reviewed by a third professional so as to identify any important limitations or weaknesses. If the court refused to appoint a more objective, neutral evaluator, then he would hire an expert who would do the job right. In any case, the postponement meant that, at a minimum, no changes in custody would take place for 6 to 9 months given the congested court calendar. In the meantime, he

urged Ms. Deer-Doe to cheer up, continue to be a wonderful mother, and to be on her very best behavior so as not to give her soon-to-be-ex-husband any ammunition in his campaign against her and motherhood.

A Bigger Problem: The Legal and Emotional Context of Custody Disputes

We could conclude our monograph here with this summary: There is essentially no psychological science to support the measures and constructs designed specifically for the assessment of child custody arrangements for individual children. Moreover, established measures of clinical constructs must be used with caution due to threats to their validity and questions about the relevance in the custody context of the constructs they assess. We also could conclude that the state of psychological science is too limited to reach clear conclusions about controversial issues such as children's wishes, overnight visits, or even PAS, and remind the reader that the burden of proof falls on proponents of a particular hypothesis or recommendation. To these three points, we could add questions about ethics and professional practice—for example, potential concerns about systematic bias on the part of evaluators, questions about whether evaluators should address the "ultimate issue" (i.e., recommend specific custody arrangements), and worries about a battle of experts when each side hires its own evaluator.

However, we believe there are bigger problems in custody evaluations than shoddy science, and we also believe that consideration of these broader issues points the way to some promising solutions for custody evaluations, children, and families. Thus, we turn now to examine the more general literature on children's adjustment to their parents' separation and divorce. After this, we outline three general recommendations that we consider in light of psychological research, legal analysis, and professional responsibilities including various issues we raised about Dr. Hagan's custody evaluation.

AVERAGE EFFECTS AND VARIATION IN THE WELL-BEING OF CHILDREN FROM DIVORCED FAMILIES

There is a large, sophisticated, multidisciplinary research literature on how children are affected by parental separation and divorce. We cannot review many original sources from this literature in this limited space, although we have done so elsewhere (Emery, 1999b). In the following section, we offer an overview of the major conclusions researchers have drawn. After this, we consider what factors predict children's more or less adequate adjustment. For present purposes, research on the average well-being of children from divorced families is of interest primarily as a starting point for examining predictions of individual differences in outcome, one of the main goals of a custody evaluation. Thus, we review this extensive literature only briefly.

On average, parental divorce is associated with an increased risk for a variety of psychological problems among children