

EMPIRICAL AND ETHICAL PROBLEMS WITH CUSTODY RECOMMENDATIONS: A Call for Clinical Humility and Judicial Vigilance

Timothy M. Tippins
Jeffrey P. Wittmann

This article proposes a four-level model of clinical inferences to analyze the psychological evaluation process in custody matters. At each level the authors summarize the status of the relevant psychological literature and conclude that, as clinicians respond to the ultimate issues (e.g., who should be the custodial parent) the empirical foundation for such conclusions is tenuous or non-existent. A jurisprudence argument is also made that such opinions should be routinely excluded from the fact-finding process. Given the significant potential for specific custody recommendations to limit personal liberties and the trajectory of a child's life, the paucity of relevant research available in this area, and profound evidentiary issues, such recommendations should be viewed as ethically inappropriate. A model for what clinicians can ethically say to courts is proposed.

INTRODUCTION

Forensic psychological assessments in contested custody matters are often pivotal documents that can have a dramatic effect on the trajectory of the litigation and, ultimately, on the form a particular child's life will take after judicial disposition. The courts afford culturally sanctioned weight to behavioral science, and these documents are often eagerly awaited because of their potential value in providing leverage for one side over the other and for their capacity to move cases toward stipulation. Indeed, recent empirical work has confirmed the high perceived value of clinicians' involvement in matrimonial matters by both attorneys and judges (Bow & Quinnell, 2004) despite earlier findings that suggested a less favorable view by legal professionals (Felner, Rowlison, Farber, Primavera, & Bishop, 1987; Melton, Weithorn, & Slobogin, 1985).

Most relevant to the current review is the finding by Bow & Quinnell (2004) that one of the components most valued by legal consumers is a specific recommendation for custody and that the overwhelming majority of judges and attorneys believe that psychologists should directly address the ultimate issue before the court. This latter belief is incongruent with the ongoing controversy about whether evaluators should make such specific recommendations, a controversy that caused the American Psychological Association itself to hedge on the issue in its *Guidelines For Child Custody Evaluations In Divorce Proceedings* (APA, 1994) and that has caused some scholars to draw the conclusion that this practice should cease altogether due to the serious limits of the psychology discipline's knowledge base on custody topics (O'Donohue & Bradley, 1999; Melton, Petrila, Poitress, & Slobogin, 1997; Tippins, 2003; Wittmann, 2003) and due to philosophical and legal arguments against ultimate issue testimony in general (e.g., Heilbrun, 2001).

This review will briefly summarize the general boundaries of our psychological knowledge regarding variables relevant to making custody recommendations. We will then

outline the jurisprudential and evidentiary problems regarding psychological opinions about custody matters, followed by an outline of what we would propose should be the parameters of ethical practice in this area. The summary will be organized initially around a four-level model for understanding clinical inferences in custody matters.

CLINICAL INFERENCE HIERARCHY

In an attempt to simplify, for discussion purposes, the very complex clinical process involved in custody assessments, we would propose the following stratification of the data and inferences garnered by clinicians in these matters:

Level I (What the clinician observes). This category includes anything that the clinician observes with his/her senses, without the addition of higher-level abstractions about the observations. Included in this category would be comments in a report such as “She hung her head low and was often tearful,” “Aaron clung to his mother’s leg throughout our session,” “There is an elevation on scale two of the MMPI-2,” and “Father would not compliment mother in any way regarding her parenthood.”

Level II (What the clinician concludes about the psychology of a parent, child, or family). This category includes inferences and higher-level abstractions about what was observed in Level I without reference to custody/best-interests constructs. Examples would include the following: “Jacqueline’s attachment to her father appears quite insecure,” “Ms. Jones has strong tendencies toward substance abuse,” “Father’s style of parenting is very authoritarian,” or “Mother’s blue mood, the child’s report of her chronic sadness, and her MMPI-2 elevations suggest the conclusion that she is moderately depressed.”

Level III (What the psychologist concludes about the implications of Level II conclusions for custody-specific variables). At this level, the clinician typically uses an even higher level of abstraction about the case, making reference to custody-specific constructs such as global “parenting capacity,” “potential psychological risks of primary custody with father,” the “fit” between a child’s needs and a parent’s “parenting capacity,” and so forth. Level II inferences about individual or family psychology (“Mother is depressed,” “There is an overly close father–son alliance against mother,” etc.) are used in Level III to make inferences specifically relevant to custody and access issues (e.g., “Multiple days away from mother, this child’s primary psychological parent, will be emotionally stressful,” “There is a poor fit between Jason’s need for parental empathy regarding his ADHD and his father’s blunt and authoritarian approach,” etc.). At this level, therefore, the psychologist is still not making overt comments about what *should* happen in a family (as in the “shoulds” of Level IV), but the inferences drawn begin to have more clear connections to the ultimate issues before the court and therefore, by definition, also begin to have greater potential impact on judicial decisions regarding important personal liberties (e.g., the right to have access to one’s children, the right to do so unencumbered by supervision, the right to make decisions regarding a child’s medical care, etc.).

Level IV (The psychologist’s conclusions about the custody-related “shoulds” in the matter). Here, the clinician takes the Level III inferences and presents prescriptive conclusions about what access schedule and other plan parameters (e.g., supervised contact, etc.) *should* be put in place by the court. Although Level II and Level III include

a certain level of value judgment on the clinician's part (e.g., "It is better for a child to not display depressive or aggressive symptoms than to display these symptoms"), it is at Level IV that the value judgments become much more overt, prescriptive, sweeping, and potentially life-changing for the litigants. There is a shift at this level from lower-level assessments of the "whats" in the matter, to taking overt positions about the "shoulds" (with the implication exuded by the "expert witness" status that behavioral science has been brought to bear on the process of achieving a specific, prescriptive recommendation about a child's postdivorce life).

THE LIMITS OF CUSTODY-RELEVANT PSYCHOLOGICAL KNOWLEDGE

Level I

At this very basic level of observation and information/reporting, mental health professionals bring specialized knowledge to the table in the form of (a) a scientific method for gathering data that should be more orderly/cautious than what would be used by a lay person and (b) clinical decision rules about what observations may have psychological significance. They arrive at the clinical room with the same sense organs as nonexperts yet can interpret what is seen and heard through the lenses of an empirical and clinical knowledge base so that early psychological hypotheses can be developed. For example, whereas a lay person may hear a dad say, "Jason and I think his mom has a problem," and make little of the comment, a clinician will note the behavior for later consideration due to the possibility of father-son alignment.

Interestingly, it is at this level that the psychologist may be able to offer some of the most useful information to the finder of fact who has to sort through the dueling distortions of family life that often characterize custody petitions. As highly trained interviewers and listeners, psychologists can set the stage of the clinical setting to maximize self-disclosure about issues that may be disputed, yet important to a best-interests decision. For example, the simple and cautious recording of best-interests-related admissions on the part of parents, or reports about family life on the part of children, can provide a valuable glimpse into how the family functions outside of the polarizing effects of litigation (e.g., hearing a father admit that he actually did strike mother in front of the children, hearing a child's tearful report of abuse at the hands of a drunken mother, listening to and recording a teenager's fervent desire to remain close to friends in a particular school system, etc.). It is the Level I observations that can assist in creating a more efficient fact-gathering process for a court and that form the basis of an *investigative* function for custody evaluators that can be quite valuable.

The limitations of our knowledge at Level I are best discussed in the context of the rich literature on clinical judgment error. Although we are not aware of any substantive research on the clinical accuracy and objectivity of custody evaluators, there is a long line of judgmental heuristics research to suggest that even at the level of basic data gathering, clinicians can be extremely prone to distortions of what they observe due to various cognitive biases, attribution effects, labeling effects, illusory correlations, flawed estimation rules, and so on (e.g., Garb, 1998; Turk & Salovey, 1988). We now also have data raising questions about the accuracy of what is actually recorded by clinicians during forensic evaluations (Lamb, Orbach, Sternberg, Hershkowitz, & Horowitz, 2000) that highlights how weighty ultimate conclusions can sometimes rest on clinical sand.

Although such judgmental distortions are now well documented, and while both clinician and lay person share these distorting tendencies, it is the clinician who makes observations

equipped with a scientific method for processing what is observed, a method that is designed to at least reduce the effects of these cognitive errors. There is evidence that when clinicians appropriately apply diagnostic strategies forcing them to consider alternative hypotheses (the strategy taught in scientist-practitioner training programs) they become more accurate in their case assessments (e.g., Tutin, 1993). We believe, therefore, that the problem in the area of custody assessment is not whether mental health professionals have evaluative knowledge or methods that supercede those of lay persons. Rather, the problem is whether this knowledge is overextended into questionable higher-level inferences and whether the scientific method is consistently brought to bear on what is observed in order to correct for ubiquitous judgment errors and to avoid having an inappropriate effect on personal liberties.

Level II

At this level the clinician begins to interpret and combine the observations made at Level I (observations, test results, record review, collateral contacts) to make inferences about the presence, absence, or severity of general psychological constructs, inferences that are made based on strands of consistent and contradictory data within the case itself. The body of knowledge that can be brought to bear here essentially represents the universe of empirically supported psychological constructs about individual and family functioning. This includes, but is not limited to, inferences about psychopathology, basic and discrete parenting skills and skills-deficits, intellectual/cognitive functioning, developmental status, developmental variables (“needs”) potentially relevant to access plans (e.g., age-differences in time perception), child-temperament variables, substance abuse tendencies, attachment constructs (e.g., “primary psychological parent”), interpersonal style, criminality, domestic violence tendencies, adequacy of parent–child boundaries, available social support, impulse control, and family-level constructs (emotional boundaries, enmeshment, etc.).

The capacity for psychological tests to validly and reliably assess a host of individual and family functioning variables is well established and, in certain areas, rivals the accuracy of medical tests (Meyer et al., 2001). Even more germane to issues related to custody and access is the empirical studies elucidating relevant aspects of the parent–child relationship as they relate to positive and negative child outcomes. For example, there is a literature on the association between certain traits/disorders in parents and adjustment problems in children (Waxler, Duggal, & Gruber, 2002), on the link between specific parenting styles and positive child outcomes (Teti & Candelaria, 2002), and on the association between substance abuse patterns and negative child outcomes (Mayes & Truman, 2002). As this line of research has advanced in sophistication and complexity away from simple notions of parents affecting children in a unidirectional manner, literature has developed around the bidirectional nature of the parent–child relationship and the ways in which children impact adults psychologically (Patterson & Fisher, 2002). There are now well-established lines of empirical evidence that allow circumscribed conclusions about secure/insecure attachment patterns (e.g., Rothbaum, Weisz, Pott, Miyake, & Morelli, 2000) and the association between certain attachment patterns and psychological adjustment problems (Rubin & Burgess, 2002).

We now have a relatively robust literature about the short and long-term symptomatic effects of parental divorce on children (e.g., Hetherington & Kelly, 2002) and about a host of factors (e.g., diminished parenting quality and socioeconomic status) that may mediate some of these effects (Hetherington & Stanley-Hagen, 2002). The negative effects on

children of parental hostility and violence are also now well documented and researchers have begun to analyze in more detail the factors that diminish or amplify the negative effects (e.g., Cummings & Davies, 1994).

This literature allows clinicians to draw circumscribed conclusions about the likely meaning of a child's symptoms ("The interview data and collateral contacts suggest the oppositional behavior on this child's part that is alleged to be the result of harsh behavior on father's part is also consistent with the symptom profile that is commonly associated with normative divorce-related adjustment challenges") and about family variables that impact on children ("The narrative descriptions of frequent angry exchanges, combined with comments by the children and with the data from police reports suggest that the relationship between these parents is very prone to overtly hostile exchanges"). In addition, clinicians can serve the court at this level by applying a rigorous hypothesis-testing approach to strands of family data so that cautious psychological conclusions can be drawn to assist the finder of fact in sorting through contradictory allegations (e.g., "Although mother admits to shyness and father asserts that she is disabled by social anxiety, the psychological testing, combined with both mother's presentation and comments by her children and her therapist do not indicate that her social anxiety is seriously limiting her interpersonal functioning."). Finally, the judicious use of psychological language to articulate the stressful circumstances that litigating families often find themselves in can assist the finder of fact in achieving a more humanized understanding of the emotional nuance behind petitions and counter-petitions.

The limitations of psychological knowledge at Level II are largely related to the limited usefulness of the data at this level to draw justifiable conclusions about custody-related Level III constructs, and the procedural dilemma of whether such data should be mentioned at all if it cannot be reliably connected to such constructs. That is, a clinician has a wealth of psychological literature and case-specific observations that can be opined about at Level II (dad is quick tempered, mom is authoritarian, dad is lacking in empathy skills, mom is irrationally suspicious, this child has a reactive temperament, etc.). Such information could easily be used, within a clinical relationship that is therapeutic or non-court related, to attempt to assist a client with their personal or family problems even if many of the inferences are only hypotheses that rest on significant clinical uncertainty. However, a forensic custody evaluation is conducted in a milieu where the words that are chosen (a) will be interpreted and used by nonclinicians and (b) may be used to truncate important human liberties for parents and, by association, for children. It therefore appears that the most ethically sensitive approach to Level II inferences is to avoid their misuse by mentioning only those that on the basis of convergent, case-specific information and the specialized knowledge base of the profession will be able to be reliably tied to Level III, custody-specific variables (i.e., to global parenting capacity, global statements about a child's psychological needs, parent-child fit, risk factors associated with different access plans, etc.). As we will assert below, this means that we must be ethically circumspect at Level II: *Despite the wealth of clinically relevant information that can be gleaned at Level II, much of it should often be withheld because of the unique socio-cultural milieu in which the forensic expert is functioning.* For example, although testing, interviews, and records support the conclusion that a parent displays serious obsessive-compulsive traits, "obsessive compulsive" should be highlighted in the forensic clinician's conclusions only if the clinician can infer, based on confirmed case data and the specialized knowledge base of the profession, that these traits can be reliably associated with custody-specific inferences (diminished parenting capacity, poor fit with a child's unique emotional needs, etc.).

We are in agreement with O'Donohue and Bradley (1999) that there are profound definitional, assessment, reliability, and validity problems associated with the interview protocols and psychological tests used in custody matters that mean that, even at Level II, many erroneous inferences are likely presented to courts on a regular basis. There is no reliably established association between traditional psychological test findings (MMPI-2, Rorschach, etc.) and different parenting styles/competencies, although we do know that test profiles predict certain states of parental turmoil and maladjustment, and studies find associations between these states and child adjustment problems (Waxler, Duggal, & Gruber, 2002). Other than a small amount of literature on associations between certain MMPI-2 profiles and an increased risk of child abuse (summarized by Pope, Butcher, & Seelen, 2000), we are unaware of any robust line of research that allows the prediction of parenting behavior through the use of traditional personality instruments. There is no custody-related, structured child interview protocol that is viewed as having acceptable psychometric properties. There are also severe threats to valid inferences from parent-child observation sessions due to the reactivity of the parties to the unique custody evaluation setting, unreliable coding systems, constricted behavioral sampling, and due to the lack of research-based guidance regarding which parent-child relationship variables can be viewed as reliably predicting different levels of child adjustment under different access plans (O'Donohue & Bradley, 1999). However, researchers and theorists are beginning to offer guidance to practitioners about clinical principles that increase the validity of inferences derived from parent-child sessions (Hynan, 2003).

One of the most significant gaps in the knowledge base that forensic clinicians bring to custody matters is related to base rates and normal distributions of various child, parent, and child-parent relationship variables (O'Donohue & Bradley, 1999). For example, a common assumption among mental health professionals is that a parent's capacity to empathize with, and emotionally understand a child is a critically important skill. However, we do not have sufficient data on how the "capacity for empathy" is distributed in the population to help a court understand whether the specific level of "empathy-deficit" in a particular parent-child dyad should be viewed as so extreme as to warrant a preference for one access plan over another. The empathy a child experiences from her father may appear greater than what she experiences from her mother, yet full normative data might show that the mother-child empathy pattern still falls within normal limits. If it does, should this be weighted against the mother? Certainly, this is more of a socio-moral decision than a clinical one. We are even lacking information about the base rates of various kinds of parenting skills in the general population, raising the same kind of challenging problem: We might note for a court that a particular father appears extremely passive with regard to limit setting, yet we have little literature to guide us in deciding whether his level of passivity should be viewed as statistically extreme enough to be relevant to best interests. The clinician is often left with personal values and clinical experience to make this decision and we believe that these sources of clinical guidance are, in most cases, wholly inappropriate as a basis for suggesting access plans that might limit certain personal liberties (i.e., by constricting parental access, constricting a child's access to a parent, etc.).

Our trial consultations have made it evident that, quite unfortunately, many forensic clinicians are willing to opine about Level II inferences (a) despite being ultimately unable to make the connection to Level III or (b) more problematically, with a willingness to state a connection to Level III that represents clinical speculation rather than empirically sound conclusions. This problem is nowhere more evident than with psychological testing: clinicians are often willing to include extended, test interpretative narratives despite being

unable to make an empirically supportable connection between the many adjectives used to describe a parent or child and custody-relevant variables (leaving posturing attorneys or judges to decide if they should view a particular description as worrisome). Many clinicians are also unaware of the literature on normative elevations on personality tests among custody litigants (e.g., Bagby, Nicholson, Buis, Radovanovic, & Fidler, 1999; Butcher, 1997). They therefore over-interpret scale elevations to the detriment of the litigant despite the fact that certain scale profiles should actually be adjusted downward because of these base rate findings (McCann et al., 2001). Finally, despite the admirable efforts by certain forensic experts to develop instruments specifically designed for custody and access issues (Bricklin, 1989, 1990a, 1990b, 1992; Ackerman & Schoendorf, 1992), most forensic reviewers agree that there are profound definitional, validity, and reliability questions that need to be resolved before these instruments should be viewed as psychometrically sound for measuring the Level III, ultimate issue-related constructs that they were designed to measure (Brodzinsky, 1993; Otto, Edens, & Barcus, 2000). However, it does appear to us that these instruments may hold great *potential* value for generating hypotheses about psychological variables at Level II (e.g., for gathering information about such issues as parenting skills, how a parent is perceived by a child, a parent's intellectual competence, etc.).

At this inferential level psychological constructs are often presented as scientifically discovered "fact" without full disclosure regarding the lack of professional consensus about such constructs. Clinicians are willing to easily include comments about "worrisome enmeshment," a child's lack of "individuation" from a parent, "primary attachment figure" and so on while not stating that there continues to be vigorous debates about the meaning and implications of such constructs. Judges and attorneys are usually not aware that many of the psychological constructs that they read about, and that they use to form strategy/opinions, are not always sturdy, enduring truths and are often not transgenerational or transcultural. For example, forty years ago a parent's homosexuality was defined as a psychiatric illness whereas now it is viewed as an "orientation." The importance of a child achieving independence and individuation from a parent essentially represents a North American value-judgment that is not made as frequently in other cultures (Rothbaum, Weisz, Pott, Miyake, & Morelli, 2000). A family viewed as enmeshed to a worrisome degree by a clinician might be viewed as evidencing a loyalty and emotional closeness that is honored in another culture. Unfortunately, few clinicians make courts aware that the constructs they use to argue for different custody plans are plastic over time, very value-laden, and can change in importance depending on one's geographic longitude and latitude.

Level III

As clinicians begin to draw inferences regarding custody-specific constructs they use language that can have substantial impact on the trajectory that a custody matter takes because such inferences, at least at a theoretical level, are more closely related to the ultimate issue before the finder of fact. A statement in a clinical report should be defined as Level III if it conveys conclusions about aspects of a child's *psychological best interests*, potentially nudging the consumer of the report in certain directions regarding access planning. Conclusions about the fit between a child's needs and a parent's skills, the global, relative parenting capacities of the two parents, custody-specific test results about "preferred parent" (Bricklin, 1984) or "custody quotients" (Ackerman & Schoendorf, 1992), and the psychological risks/benefits of different custody plans all have more

perceived relevance than Level I and Level II to the custody and access decisions before the court (insofar as these constructs begin to limit the field of psychologically preferable access plans for the consumer of the report). However, such statements are still in the realm of “whats” because they ostensibly are still conveying conclusions about the *psychological state-of-affairs* of the evaluated family without specifically stating what the court “*should*” do.

The American Psychological Association, in its guidelines for custody evaluations (APA, 1994), begins to suggest a nomenclature for Level III in its conclusion that the “fit” between a child’s psychological functioning and developmental needs and a parent’s functional ability to meet these needs should be the focus of custody assessments. It is at this level where the clinician begins to make a leap from Level II inferences regarding such matters as security of attachment, parenting skills, developmental needs, and so on to opinion statements about the goodness of fit between what was concluded about the child at Level II (psychological functioning and developmental needs) and what was concluded about the parents at Level II (functional skills of parents). It is, therefore, at this level of inference that the clinician begins to have a more potentially substantial impact on the ultimate decision (and, therefore, we would argue that the ethical mandate for the conclusions to be empirically sound and grounded in specialized knowledge of the profession becomes substantially more paramount).

We would argue that the only ethical statements that can be made by clinicians at Level III include a summary for the court of notable psychological risks and/or advantages associated with various access plans, but that such statements can only be made with clearly articulated qualifications, cautionary statements to the court, and references to the limitations of the evaluation methods used. It will be argued below that the statements that clinicians can make at this level that are within acceptable ethical and scientific parameters are extremely limited. However, we also believe that “helpfulness” to the finder of fact (Melton et al., 1997) as a guiding principle suggests that child-focused, yet constricted, statements about potential risks/advantages, as long as they are grounded in case-specific facts and reliable empirical literature, represent a forensic work product that is ethical, useful to the court, and potentially valuable to both the child in question and society at large. Examples of the kinds of constricted, risk-focused conclusions include the following: (1) “Primary placement with the father in this matter includes the risk of long-term exposure to Mr. Smith’s chronic tendencies toward derogatory and violent behavior with women.” (2) “Primary placement with mother includes the risk of substantial, weekly exposure to her now, well established and self-admitted tendencies toward alcohol abuse while under stress, tendencies that interrupt her ability to meet this toddler’s needs for careful supervision.” (3) “Jonathan has been diagnosed with severe ADD and a mild psychotic disorder that all clinicians agree interferes with his academic and social functioning—a notable risk associated with primary placement with his father, therefore, involves the fact that Mr. Smith is openly resistant to giving his son the medications that appear to improve his emotional adaptation.” (4) “A risk factor associated with primary placement with mother includes her chronic inability to set firm and predictable limits on her eight-year-old son’s behavior, as indicated by her behavior in session, the record review, and consultation with the child’s therapist—an inability that is especially relevant given the fact that Jason has begun to engage in delinquent behavior in the community. Placement with father appears to hold the advantage of firm and predictable limit setting for Jonathan, a parental pattern that has been established in the research literature to have a positive association with better child outcomes than more passive approaches.”

Such a risk-focused approach does not take an affirmative stance about whether the specific risks should be determinative and allows substantial room for negotiation, as well as for the finder of fact to make the ultimate, socio-moral value judgments about whether such risks should be attended to in the final access plan and how they should be weighted. It also rests on an implicit value judgment that we should make explicit: the value to children and families of minimal invasiveness by mental health professionals except where case-specific and research-based data can reliably suggest important, child-relevant risks/advantages.

Unfortunately, with the exception of such constricted statements of risk/advantage, we conclude that there is little that can be reliably said by mental health professionals at Level III. The concept of the adequacy of “fit” between a child’s needs and a parent’s capacity has not been operationalized by the psychology discipline; that is, individual clinicians are left to their own idiosyncratic definitions of what adequate “fit” means. There is a quite paltry, yet growing, empirical literature on reciprocal effects between parents and children that *might* be relevant to custody (Patterson & Fisher, 2002) but no one to our knowledge has tested these constructs across different access plans. There is no established metric for assessing the goodness of parent–child fit and there has been no attention in the literature to the ways in which parent–child mismatches might be associated with eventual, positive child outcomes (a mother’s superior capacity for emotional attunement may align nicely with the mental health discipline’s emphasis on empathy and compassion, whereas the father-CEO’s high academic expectations may produce other valuable achievement-focused outcomes for a child that are less valued by the same discipline).

Although the substantially increased complexity of our understanding of child and family functioning has caused the somewhat simplistic notion of “primary psychological parent” (Goldstein, Freud, & Solnit, 1973) to be less emphasized in recent years, we continue to find it present as an operative construct in a fair number of custody assessments. It is important to emphasize that, at least as a psychological construct that should be determinative for access planning, this construct is hotly debated and very controversial; recent reviewers have suggested that it may be time to set it aside in favor of more complex and multidimensional views of children’s attachment needs (Kelly & Lamb, 2000). As such, it has little professional consensus as a determinative variable and its use as a centerpiece by a clinician for forming a rationale for a particular custody plan should be viewed as highly suspect (see Level II comments above).

Other custody-specific constructs often used to argue for particular access plans include the “parent of choice” construct presented by Bricklin (1984) and the “parental custody index” offered by Ackerman and Schoendorf (1992). Both systems offer a welcome attempt to operationalize custody-relevant psychology variables and to standardize assessment approaches. Unfortunately, as noted above, there are numerous substantive questions about the validity and reliability of these instruments and, even more apropos to the current discussion, a lack of consensus among mental health professionals regarding how such instruments should be validated (Brodzinsky, 1993). Until such a consensus is reached we believe that the use of such constructs in forensic reports, especially at Levels III and IV, implies a level of scientific consensus and certainty that is inappropriate. In addition, we believe that a more procedurally appropriate approach to developing custody-related instruments is to focus research on better assessment techniques for Levels I and II, and that trying to “measure” such socio-moral constructs as “best interests” or “best custodian” (the Level III constructs and the “shoulds” at Level IV) should be deemphasized because they represent, in a sense, an attempt to measure the incalculable (partially

because such value-laden constructs can reasonably be given different definitional contours by legal and mental health professionals who have different preferences/values, yielding little consensual agreement about what is important to a child and to the definition of the construct).

Another challenge with Level III inferences is the fact that the vast majority of psychological research that informs Level II inferences was not designed specifically to inform those at Level III. For example, while the developmental literature might allow a clinician to draw reasonable conclusions about attachment patterns in a particular family and about potential risks to healthy attachment associated with different custody schedules, there is virtually no empirical literature that allows the clinician to then, on the basis of this literature, reliably suggest specific access plans. This is because we do not know how the effect of such patterns changes in the context of the enormously complex emotional life of the separating family. Long-held truisms about attachment may morph or become invalidated under particular schedules, with particular children and parents who display a unique array of psychological attributes.

It is our opinion that the most egregious manner in which Level III custody inferences are used in forensic reports involves the lack of open admission to the court of the degree to which such inferences rest on substantial uncertainty and on the subjective values of the individual clinician or the mental health profession at large. Informed clinicians can take reasoned positions against a wide variety of custody-relevant assumptions (e.g., O'Donohue & Bradley, 1999; Melton et al., 1997) including: (1) continuity in social relationships should be paramount (What about the capacity of a child to be strengthened and enriched by change?); (2) continuity in the relationship with the primary psychological parent, often assumed to be the mother in this culture, should be paramount (because there is a rich literature regarding the benefits of father involvement that may argue against this position); (3) children should never be placed with a parent who has engaged in domestic violence (but occasionally domestic violence is circumscribed to separation-related events and the nonviolent parent may have other meaningful forms of dysfunction such as alcoholism that also place children at risk); and (4) siblings should be kept together (yet occasionally children have very distinct and contradictory personal needs that may suggest the value of dividing the siblings). These represent only a small sampling of the highly subjective value choices that often lay behind the positions clinicians take at Level III and that are often opined about in forensic reports without adequate disclosure of professional uncertainty.

The problem here is not that clinicians have nothing to say about these matters (we know about the effects of domestic violence, alcoholism, etc. on children). However, we have virtually no empirical literature to guide us in the selection of access plans among competing, complex clinical realities (i.e., in the weighing of different detriments and advantages). Despite this state of affairs, it is at Level III that clinicians often present their logic and personal values under the guise of behavioral-scientific truth without disclosing that we have no reliable clinical method to do this weighing. Simple logic and subjective values do not represent specialized knowledge.

Finally, it is not uncommon for clinicians to use a child's current symptomatic response to a family plan as one argument for an alternate access plan. However, while it is true that substantial data is accumulating on the normative effects of divorce on children (e.g., Hetherington & Kelly, 2002), we have no empirically proven method for teasing out whether a child's symptoms are best viewed as a time-limited response to a court dispute (that will subside postdisposition regardless of the access plan) or a direct response to a specific custodial schedule.

Level IV

As noted above, it is the specific custody recommendation that is most valued by the consumers of forensic evaluations (Bow & Quinnell, 2004). At Level IV, the clinician begins to communicate a substantial narrowing of the preferred custodial plans for a particular child and, quite often, specifically chooses a single plan as the most beneficial or least detrimental (right down to the hours of the day or night a child should be with each parent and whether or not there should be sole or joint legal custody). The professional posture of the forensic clinician who chooses to communicate at this level shifts dramatically from having described for the court the “whats” of a family (Levels I, II, and III) to a specific, prescriptive recommendation regarding what a court “*should*” do.

We are in agreement with forensic scholars who conclude that clinicians should routinely avoid addressing the ultimate issue before the court in the form of a specific recommended access plan (e.g., Karras & Berry, 1985; Melton et al., 1997). Drawing conclusions about the ultimate issue before the finder of fact blurs critically important boundaries between the person invested with the power to make socio-moral and social-control decisions (the judge) and the expert witness who is hired to assist the court (e.g., the psychologist). This argument holds that, even if one were to develop a family blood test that could perfectly predict the custody plan that would lead to the best psychological functioning on the part of a child, clinicians should still refrain from making that specific recommendation (although they might be able to say many things at Levels I, II, and III to “tip their hand”). A clinician might be able to say that “the family custody blood test” predicts with 99% certainty that primary custody with mother would lead to less behavioral disturbance on the part of the girls in this family. However, specifically saying “therefore, the court should place the children with their mother” would still arguably be inappropriate. We are not prepared to assert, and it is beyond the purview of this article to consider, that the policy judgment of Federal Rule of Evidence 704 and its common law parallels *generally* permitting ultimate issue testimony should be repudiated. However, it is imperative that mental health witnesses carefully circumscribe their testimony so as not to blur the critical distinction between the function of an expert witness to expound from the specialized knowledge base (i.e., the “whats”) and that of the judge to make the ultimate legal-socio-moral determination (i.e., the “shoulds”).

However, even if one asserts that experts should be able to address the ultimate issue as a general proposition, our review of the literature suggests that we are quite far from the development of a “family custody blood test” and that the likelihood of developing predictive power that would warrant specific custody recommendations is extremely remote. Yet an extremely common part of forensic assessments, and often the part of the report read first by lawyers and judges, is the section in which clinicians dare to tell the court what days and times the children should be with what parent.

The magnetic forces of trying to make private practices survive and of having legal advocates and judges tug at clinicians for very specific direction make the regular choice to provide specific custody recommendations understandable, yet no less egregious. Any clinician who recalls training in multivariate statistics and prediction will, if honest, attest to the astronomically complex cognitive task of making relatively circumscribed behavioral predictions, let alone combining and weighing the many inferences drawn at Level II and III into a specifically recommended access plan. However, there is not one piece of research that we know of that supports the notion that clinicians can reliably and validly engage in this multidimensional inferential process in a way that yields recommendations that are

correlated with positive child adjustment. There is no psychological technology existing for weighing and combining the complex factors assessed at Levels I, II, and III into access plan recommendations that are empirically established to yield better child adaptation than alternative plans. Theoretically, at Level IV, the clinician has somehow combined and weighed intra-psychic, behavioral, dyadic, family-level, extended family-level, and cultural issues into a direction for the court that is presented as representative of behavioral science, and this occurs despite the complete lack of evidence that we are able to do so reliably and validly.

There is a growing literature on the correlations between various custody plans (sole vs. joint) and positive/negative adjustment in children (e.g., Bauserman, 2002). Assuming that this literature grows and becomes more robust, it may serve as a useful source of knowledge for judges and policy makers. However, given the scores of variables that are idiosyncratic to specific families, it may well be impossible to generate empirical prediction data that could be used to back up specific custody recommendations for specific families (Melton et al., 1997). We will paraphrase and extend the phrase offered by O'Donohue and Bradley (1999) in the following fashion: when clinicians make a custody recommendation they are essentially portraying themselves as knowing the best plan for a child after answering the question "*What initial state of mother, father, children, extended family, and other important relationships, and what contextual variables, when considered in the context of the scores of potential access plans for a given family, will result in what later negative or positive state in the children?*" To say that the mental health profession is inadequately prepared to make such multivariate decisions is an understatement, yet day in and day out forensic reports imply that we are quite able to do so.

Another way of stating the dilemma at this level is the following: there is no evidence in the empirical literature that current interview protocols, traditional psychological tests, or custody-specific tests are in any way able to reliably predict child adjustment to different access plans, yet 94% of evaluating psychologists still make such recommendations (Bow & Quinnell, 2004).

At this level, as in Level II and III, there is often a long list of behind-the-scenes values on the part of the clinician that is guiding their choice of data, how they interpret it, and what they recommend. Unfortunately, many clinicians make specific custody recommendations without explicating their often subjective beliefs about what is "good" and "bad" for children. It should be emphasized that the purpose of behavioral science is to empirically guard against the excessive encroachment of subjective values, or at least to guard against unwarranted assumptions. Consequently, clinicians arrive at court cloaked in the presumed objectivity of behavioral science when many subjective values are directing their conclusions (e.g., boys do best with their fathers, children do best with their primary attachment figure, joint plans are bad for young children, etc.). Few clinicians are up front with the legal consumers of their reports about the degree to which the assumptive underpinnings of their recommendations are highly subjective and controversial. In addition, the often emotionally infused desire to *advocate* for a particular child's interests (rather than for the data) can itself introduce profound and distorting bias into the clinical process (Wittmann, 1985).

Given the virtual nonexistence of empirical data supporting the predictive validity of psychological interviews and testing with regard to specific access plans, ethical issues become quite problematic. The ethical principals of psychologists and Code of Conduct (APA, 2003) state the following in Ethical Standard 2.04:

Psychologists' work is based upon established scientific and professional knowledge of the discipline.

However, we have no empirical knowledge/support for our ability to make specific custodial predictions for specific families. Ethical Standard 9.02(a) states the following:

Psychologists administer, adapt, score, interpret, or use assessment techniques, interviews, tests, or instruments in a manner and for purposes that are appropriate in light of the research on or evidence of the usefulness of the proper application of the techniques.

Unfortunately, because there are no clinical techniques currently available which have been shown to reliably predict positive or negative child adaptation under different access plans, it would seem that any specific custody recommendations based on current instruments and techniques would be ethically inappropriate.

Ethical Standard 9.02(b) states the following:

Psychologists use assessment instruments whose validity and reliability have been established for use with members of the population tested. When such validity or reliability has not been established psychologists describe the strengths and limitations of test results and interpretation.

While there are many assessment techniques with established reliability and validity for assessing Level I and Level II variables, the lack of similarly valid and reliable instruments for Levels III and IV (i.e., for predicting adjustment under various plans) is an ethical indictment of any specific custody recommendation.

Ethical Standard 9.06 states the following:

When interpreting assessment results . . . psychologists take into account the purpose of the assessment . . . They indicate any significant limitations of their interpretations.

Again, given the paucity of research noted for Level IV, psychologists are clearly required to inform custody litigants of the lack of validity and reliability of the interview and test techniques they use if they intend to provide specific recommendations.

Finally, Principal A states the following:

Psychologists strive to benefit those with whom they work and take care to do no harm.

We assert that specific custody recommendations, given their lack of adequate empirical foundation, can hold significant potential for harming a child because of the dramatic effect that such recommendations can have on the individual, familial, academic, and social trajectory of that child's life (given the absence of a knowledge base to reassure the finder of fact that such a "trajectory" has been shown to lead to reliable, positive child outcomes).

Ethical considerations evolve as a discipline becomes more self-aware. To date, most clinicians have offered specific recommendations in a well-intentioned and sincere effort to provide guidance that will help children vulnerable to the stressors of family reorganization. In addition, they have been taught by the legal profession that their work product is highly valued, reinforcing this professional behavior. However, given the various ethical dictates noted above, and our more evolved self-understanding as a discipline regarding the limits

of our knowledge, it is our opinion that the time has come for specific custody recommendations to be viewed as unethical for at least the foreseeable future.

CUSTODY RECOMMENDATIONS: AN EVIDENTIARY ANALYSIS

Given the enormous impact that custody evaluations can have on the lives of litigants and their children, and given the remarkable limitations on the underlying professional knowledge base and lack of consensus as to uniform methodology, it is important to examine these opinions within the framework of the law of evidence that governs admissibility. Notwithstanding the increased attention to the gatekeeper responsibility of the trial court sparked by the U.S. Supreme Court decision in *Daubert v. Merrell Dow Pharmaceuticals* (1993), custody opinions by mental health witnesses have largely escaped the serious judicial scrutiny of reliability and relevance that is so central to the fundamental question of admissibility. As one scholar has trenchantly noted, "It is striking that one of the most important categories of cases, the future of our children, has ignored the call for trial judges to address threshold scrutiny of the reliability and relevance of expert testimony" (Shuman, 2002, p. 139).

Opinion evidence, whether from lay witnesses or experts, has long been a vexing and multi-dimensional issue. Writing of opinion evidence generally, McCormick, at § 12, states: "In the outer circle of collateral fact near the rim of relevancy, evidence will be received with relative freedom; but as we come closer to the hub of the issue, the courts are more careful to call for details instead of inferences."

Juxtaposition of McCormick's statement with the four-level analysis presented above is graphically depicted in Figure 1.

Only Levels II, III, and IV are levels at which inferences are being drawn. At Level I, the clinician is reporting "factual" observations made during the clinical interview. While his or her investigative expertise and interviewing skill may well have uncovered the observed fact and the observation may strike right at the core of the best interest issue, his or her testimonial status in reporting the observation is nonetheless that of a *fact witness*. For example, if during the interview, the father became angry and struck the child physically, his conduct would impact upon the best interest issue quite directly. Yet, from an evidentiary perspective, the clinician's secretary would be quite as competent as the evaluator to testify to the occurrence, assuming he or she observed it, because its relevance does not depend upon any inference being drawn by the witness.

Once we move to the *inferential levels*, however, one can readily observe that the closer to the core issue, i.e., the centermost circle, the opinion strikes, the greater its potential impact on the outcome of the case. Accordingly, under McCormick's formulation the greater the judicial vigilance and evidentiary rigor must be. Similarly, it may be said that the more questionable the reliability and validity of the premises underlying the opinion, that is, the principles, methodology, and techniques of the profession whose expertise is being proffered, the more muscular the court must be in exerting its prerogatives as evidentiary gatekeeper.

In terms of the multilayered paradigm posited above, as the witness proceeds from Level I to Level IV, from concrete factual observations to the higher levels of abstraction, the opinion strikes closer and closer to the heart of the case and at the same time becomes weaker and weaker in terms of its scientific validity and evidentiary reliability.

In this regard, McCormick makes an interesting statement:

Undoubtedly there is a kind of statement by the witness which amounts to no more than an expression of his general belief as to how the case should be decided or as to the amount of

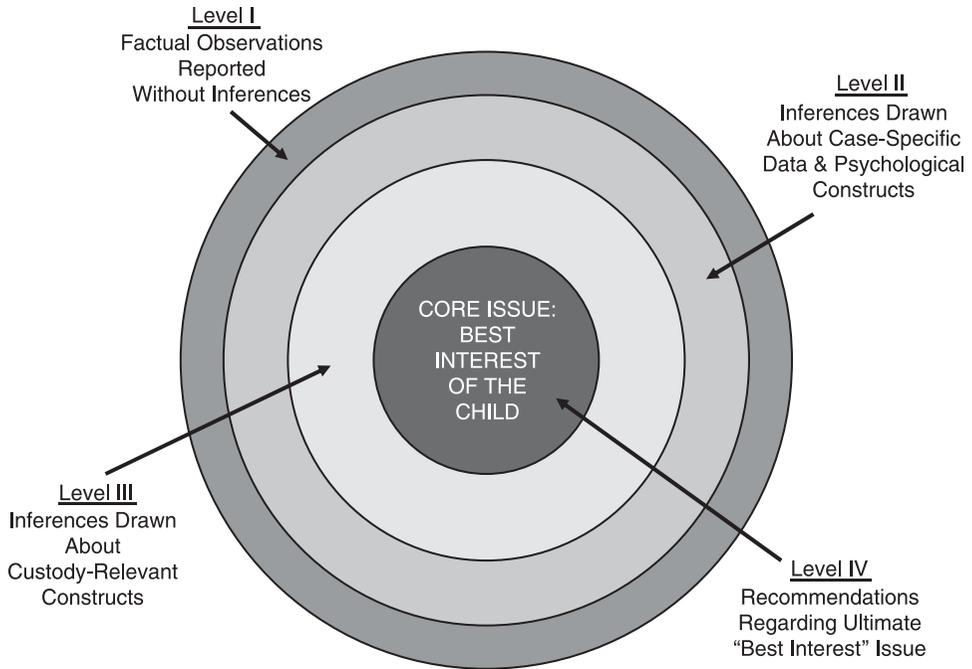


Figure 1

damages which would be just. All courts exclude such extreme, conclusory expressions. There is no necessity for this kind of evidence; *to receive it tends to suggest that the judge and jury may shift responsibility for the decision to the witnesses*. In any event, it is worthless to the trier of fact [*italics added*] (McCormick, § 12).

Interestingly, the specific custody recommendations under discussion fall precisely into this category of “extreme, conclusory expressions” of opinion as to “how the case should be decided.” It is particularly troublesome because, as noted above, such recommendations are not anchored to sufficient empirical research. Exacerbating the issue further, the very overreliance upon the opinion of the witness, noted by McCormick and cautioned against even in some high judicial quarters,¹ is exactly what often comes to pass with respect to custody evaluations. As one commentator observes,

The use of court-appointed experts whose opinions determine the outcome of many, if not most, contested custody cases effectively delegates judicial power without formal legislative approval (Shuman, 2002, p. 161).

When that happens, the court may be basing its decision on personal value judgments of witnesses who happen to have professional credentials. However, those credentials do not entitle the witness’s personal, as opposed to professional or scientific, judgments to be admitted, let alone carry weight. When an expert witness “does not testify on the basis of the collective view of his scientific discipline” and where “no understandable scientific basis is stated,” “personal opinion, not science, is testifying. . . .” (Daubert, 1995; Turpin, 1992). Cogently stated, “something doesn’t become ‘scientific knowledge’ just because it’s uttered by a scientist. . . .” (Daubert, 1995).

To the extent such personal opinions are admitted under the guise of science and are allowed to impact the outcome of the case, the process is tainted and judicial power is usurped. Therefore, when confronting the proffer of such opinions, the court, in its role as gatekeeper, has every reason to be Draconian in its application of the rules of evidence and traditional common law safeguards.

Evidence doctrine is well suited to the task of safeguarding the judicial process from pseudoscientific opinions. Indeed, the *raison-d'être* of its many rules is the exclusion of unworthy proof. If the law of evidence were concerned only with allowing evidence in, it could be stated in a single sentence: "Evidence is admissible when it is relevant"² (*People v. Scarola*, 1988). If that were the only thing to be said on the subject, such legal scholars as Wigmore, McCormick, and others could have saved immeasurable hours and intellectual energy, not to mention trees. Significantly, the vast body of evidence law is aimed at keeping unworthy evidence from tainting the mind of the fact finder.³

ASSESSING RELIABILITY: THE DAUBERT STANDARD

In 1993, the U.S. Supreme Court, in *Daubert*, established a new standard for assessing the admissibility of expert scientific testimony. The Court subsequently held that the new standard is to be applied to the assessment of all expert testimony, whether grounded in science or some other species of specialized knowledge (*Kumho*, 1999).

In establishing a new standard, the Court explicitly rejected the so-called *Frye* rule, under which, the trial court, instead of directly determining the validity of the science underlying the expert's opinion, looked to whether the principle or method had gained "general acceptance" within the relevant scientific community.

Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs (*Frye*, 1923, p. 47).

Under *Frye*, the court engaged in "counting scientists' votes, rather than on verifying the soundness of a scientific conclusion" (*People v. Wesley*, 1994; *Jones v. United States*, 1988; *State v. Montalbo*, 1992).

It has been noted that "*Daubert* could not have come at a better time" because of the proliferation of harsh criticism of expert testimony that is predicated upon "junk science," with some of the harshest criticism being directed toward the testimony of mental health witnesses (Lyon & Koehler, 1996; Foster & Huber, 1997).

Daubert is often cited as authority for the declaration that the trial judge is the "gatekeeper," who is charged with meaningful scrutiny of proffered scientific evidence (Reagan, 1999). In fact, judges have always served in that role. Not only with respect to expert testimony, but, indeed, with respect to all evidence, the common law trial judge sits as referee, calling balls and strikes in accordance with the law of evidence, deciding what proffered evidence is admissible and what is not. The distinct contribution made by *Daubert* was that it *reclaimed* this vital role that the *Frye* test essentially delegated to extra-judicial venues.

Yet, the judicial mission under either *Frye* or *Daubert* is the same. It is to ensure that only opinions grounded in valid scientific principle and methods make it across the

drawbridge, and that those that only appear to be science, that is, pseudoscientific assertions, do not (*People v. Wesley*, 1994). The difference is that under *Frye* the court effectively *delegates* this determination to the relevant professional community, accepting as a given that the science is “good science” where it has gained general acceptance within that community. Under *Daubert*, the court *directly* assesses whether proffered testimony is so grounded, upon consideration of scientifically relevant criteria, such as whether the theory or technique can be or has been tested, whether it has been subjected to the rigors of peer review and publication, how precise it is as measured by the known error rate, as well as whether it is generally accepted in the relevant scientific community (*Daubert*, 1993).

Although, under *Daubert*, the applicable criteria of reliability may vary from issue to issue depending upon what subject matter is at hand (*Kumho*, 1999), the *standard* it enunciated is applicable to *all* expert testimony and was sufficiently clear that it was later incorporated into an amended Rule 702 of the Federal Rules of Evidence:

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

It is because specific custody recommendations are *not* based on sufficient empirical data and therefore, cannot currently be “the product of reliable principles and methods,” that they ought not to be admissible.

GENERAL ACCEPTANCE—DEFINING THE RELEVANT SCIENTIFIC COMMUNITY

Frye and *Daubert* are both federal court decisions. While some states adopted the *Daubert* standard, others continue to adhere to the *Frye* test of general acceptance (*People v. Wesley*, 1994). Even under *Daubert*, however, *one* of the criteria for determining reliability is to look at whether the principle or method in question has gained general acceptance in the relevant scientific community.

Both *Frye* and *Daubert* establish “a test of admissibility,” the purpose of which is to ensure that the evidence presented is *reliable* (*People v. Wernick*, 1996). The distinction between the two is simply that *Daubert* undertakes to do so directly. However, when determining admissibility under *Frye*, the court does not *directly* assess the scientific validity, that is, the evidentiary reliability of the proffered testimony. Rather, as noted earlier, it looks to whether the scientific principle or procedure at hand has gained general acceptance “in the particular field in which it belongs” (*Frye*, 1923, p. 47).

Under *Frye*, the court exclusively “looks to general acceptance of the procedures and methodology as reliable within the scientific community,” whereas under *Daubert* general acceptance is simply *one* of the factors considered (*People v. Wernick*, 1996). Under either standard, proper definition of the relevant scientific community is important. For a *Frye* court, because general acceptance is the *only* factor considered, proper definition is nothing less than critical.

Defining the relevant scientific community with respect to the reliability of specific custody recommendations is not as simple a proposition as it may first seem. Because custody evaluations usually lie within the realm of psychology, certainly the evidence of reliability

ought to come from its domain—but from which quarter? Here, one must take note of the schism that exists within the field between those who anchor their conclusions upon empiricism and those who proceed by intuition.

Even more deep-seated than the division between scientists and practitioners in clinical psychology has been the split between *empiricists* and *romantics* (a terminology suggested by psychiatrist Paul McHugh of Johns Hopkins University). According to scholars of intellectual history, two broad trends can be discerned in European and American thought over the past few centuries. One trend (which McHugh calls “empiricist” but historians might call “enlightenment”) holds that reason, objectivity, and empirically verified evidence provide the surest road to knowledge. The other trend (which both McHugh and the historians call “romantic”) teaches that intuition, empathy, and subjective insights can provide deeper and more authentic understanding than can mere reason⁴ (Wood, Nezworski, Lilienfeld, & Garb, 2003, p. 93).

Notwithstanding the fact that mental health witnesses sometimes attempt to escape the clutches of cross-examination by claiming psychology is not a science, but an art, or a hybrid of art and science (Gould, 1998, p. 53), a claim that carries many other evidentiary implications, none of which operate in favor of admissibility, the mainstream view or consensus within the profession is to the contrary. As one historian of the mental health profession states:

As a result, empiricism has become the dominant perspective of contemporary psychology, gaining almost universal acceptance. There seems to be widespread agreement that scientific advances are optimally produced and conveyed under the procedures of empirical verification; other forms of inquiry do not appear to offer the compelling attraction of empiricism (Brennan, 2003, p. 332).

Accordingly, it is both appropriate and necessary to hold the mental health witness accountable for the application of empirically supportable principles and methods. The courts have both the right and the obligation to insist that the experts whose opinions can change lives in dramatic fashion support each and every one of their inferences with specific empirical evidence. In other words, the legal system must demand that the premises and reasoning of the expert be scientifically valid.

The deep schism between the scientists of the profession and the romantics ought to be of critical concern to the legal profession when assessing admissibility. When a court assesses the general acceptance issue, under either *Frye* or *Daubert*, it must look to those within the profession who adhere to the scientific method and who insist upon empirical verification of their premises and conclusions.

Unfortunately, “scientifically minded clinicians are becoming a rapidly dwindling minority within their profession” (Lilienfeld, Lynn, & Lohr, 2003, p. xi). Indeed, the term “schism” may be too mild to describe the phenomenon. “Today, however, calling it a ‘gap’ is like saying there is an Israeli-Arab ‘gap’ in the Middle East. It is a war, involving deeply held beliefs, political passions, views of human nature and the nature of knowledge, and—as all wars ultimately involve—money, territory, and livelihoods” (Lilienfeld et al., 2003, p. xiv).

The implications of this dwindling attachment to scientific method with respect to therapeutic settings aside, in the forensic context such unanchored opinions can determine whether or how often a specific child gets to see a particular parent. Here, with substantial individual rights hanging in the balance, it clearly becomes the responsibility of the courts

to keep unscientific opinions regarding access plans out of the process entirely. Whatever epistemological merit there might be to the romantic position, the courts, in assessing evidentiary reliability, must look to the empiricists for their assessment of the existing research data that supports or disconfirms tendered theories.

It is therefore critical that the courts define the relevant community with care because that definition can determine the outcome of the general acceptance issue. If, for example, the *Frye* court had defined the relevant community as polygraphists, rather than looking to “physiological and psychological authorities” (*Frye*, 1923, p. 54), the polygraph likely would have been found admissible. Likewise, opinions based on astrology would be admissible were the relevant community defined as astrologers and inferences drawn from the Ouija board would come in if we look only to those who believe it has predictive validity. The Supreme Court acknowledged as much in *Kumho*, finding the general acceptance factor of little assistance “where the discipline itself lacks reliability, as, for example, do theories grounded in any so-called generally accepted principles of astrology or necromancy” (*Kumho*, 1999, p. 1175).

Likewise, the reliability of custody evaluations and specific recommendations cannot be accurately assessed by reliance upon those in the clinical community who have largely departed the realm of science. The court, for an accurate assessment of the state of specialized knowledge, must turn to psychologists who emphasize science and research, and empiricists, who recognize that intuition and subjective insight are inappropriate bases for forensic conclusions and who bind themselves to the scientific method. To do otherwise is akin to looking to astrologers or necromancers to assess the scientific reliability and validity of star-based predictions and séances.

A court looking to general acceptance either as a standard or as one of several criteria of reliability is tracking for the presence or absence of controversy, a point made by Chief Judge Kaye in her very instructive concurrence in *People v. Wesley* (1994):

The point of noting controversy about the reliability of the forensic technique is not for our Court to determine whether the method was or was not reliable in 1988, but whether there was consensus in the *scientific* [italics added] community as to its reliability. The *Frye* test emphasizes “counting scientists’ votes, rather than on verifying the soundness of a scientific conclusion.” (citations omitted) Where controversy rages, a court may conclude that no consensus has been reached (p. 439).

Two points are worth making in this regard. First, as the writings of the several scholars and commentators referenced herein, as well as others not cited, make clear, controversy rages with respect to the validity of specific custody recommendations. Somewhat curiously, it is a rather one-sided controversy in that much has been published in the peer-reviewed literature delineating the absence of empirical support for specific recommendations, while scholarly argument supporting the empirical foundations for such recommendations is scant to nonexistent even though the practice of rendering recommendations has become commonplace.

Second, in determining the admissibility of such recommendations, it is clearly the opinion of the *scientific* community as to general acceptance that is required. That is the essence of the *Frye* test. It is also why the opinions of nonscientific practitioners will not be helpful and why the courts should turn only to carefully chosen scientist-practitioners and, more often, to research psychologists. It is this latter group of professionals who have been trained from the start to understand that replicated behavioral research in conjunction with cautious hypothesis testing is the only way to guard against subjective bias, idiosyncratic

and unverified theory, and the excessive encroachment of personal values into custody-related inferences that can change the trajectory of a child's life (Levels III and IV).

Looking to scientist-practitioners and researchers, rather than those who rely on intuition and theory that is without a solid research base, serves another purpose as well. New York's highest court has noted that courts ought to be cognizant of the potential economic interest of those to whom it looks for evidence of general acceptance: "A *Frye* court should be particularly cautious when—as here—"the supporting research is conducted by someone with a professional or commercial interest in the technique" (citation omitted) (*People v. Wesley*, 1994, p. 464)" (Kaye, C. J., concurring). Thus, the question of reliability ought not to be submitted to the custody evaluators who have an economic interest in the admissibility of their opinions.

BURDEN OF PROOF

Finally, when it comes to establishing that there is a valid scientific basis for the opinion of the expert, the burden of proof rightly falls upon the proponent of the evidence. "A basic tenet of science is that the burden of proof always falls squarely on the claimant, not the critic. Consequently, it is up to the proponents of these techniques to demonstrate that they work, not up to the critics of these techniques to demonstrate the converse" (Lilienfeld et al., 2003, p. 3; *Tyson v. Keane*, 1998; *People v. Wesley*, 1994).

QUALIFICATION: A QUESTION-BY-QUESTION ISSUE

The first step in taking the direct testimony of an expert is to have the witness accepted by the court as "qualified as an expert." All too often, once the witness has been so "accepted," what ensues is an intellectual imbroglio, with the witness being allowed to opine on matters that do not fall within the range of his or her expertise.

When looking to the issue of witness qualification, a distinction must be made between the question of whether the witness is generally qualified to speak for his or her profession and the discrete, but often overlapping question of whether that profession has anything relevant to say about the issue before the court that is grounded on scientifically valid principles as established by empirical research. As one commentator has noted:

By virtue of their qualifications alone, experts do not provide any assurance that their opinions rest on reliable methods and procedures. Instead, relying on experts without testing the reliability of their methods and procedures cloaks experts' value judgments under the veil of science and risks that their personal and professional characteristics bias the evaluation and the importance of information learned (Shuman, 2002, p. 160).

Significantly, under Rule 702, the witness must be qualified as having *scientific, technical, or specialized knowledge* with respect to those issues that the court determined will be helpful to the trier of fact. Once so qualified, the witness is allowed to testify in the form of opinion or otherwise, but only to the extent that "(1) the testimony is based upon sufficient facts or data, (2) *the testimony is the product of reliable principles and methods*, [italics added] and (3) the witness has applied the principles and methods reliably to the facts of the case."

This implicates Wigmore's delineation that the expert witness's testimonial capacity "in every case is a relative one, i.e., relative to the topic about which the person is asked to

make his statement. The object is to be sure that the question to the witness will be answered by a person who is fitted to answer it" (Wigmore, 1979, § 555, p. 750).

Further, Wigmore states that: "Since experiential capacity is always relative to the matter at hand, the witness may, *from question to question*, [italics added] enter or leave the class of persons fitted to answer, and the distinction depends on the kind of subject primarily, not on the kind of person" (Wigmore, 1979, § 555, p. 750).

Assuming, therefore, that the witness is generally qualified as an expert in the particular scientific discipline, the related question is whether that discipline has something to say in response to each particular question posed, a "something" that can be proven to be scientifically valid by the empirical research that constitutes the specialized knowledge base of that discipline. It was precisely the lack of such valid underlying methodology that was the subject of comment by the Supreme Court in *Kumho* where the Court sustained the lower court's rejection of expert testimony. The District Court had found the witness qualified as a mechanical engineer, yet rejected the proffered testimony because "despite those qualifications . . . the methodology employed by the expert in analyzing the data . . . and the scientific basis, if any, for such an analysis" was not shown to be reliable (*Kumho*, 1999, pp. 1176–1177; Shuman, 2002, p. 139).

In the absence of empirical research that establishes that there are reliable principles and methods at the disposal of the expert's profession, then the testimony falls outside the domain of professional opinion and into the realm of personal value judgment and theoretical assumptions. This distinction between expert-based conclusions and personal opinion is implicit in the requirement that the expert's opinion be stated in terms of "reasonable *professional certainty*" and is of critical importance in assessing the admissibility of the opinions offered in custody evaluations:

There is an important difference between an expert opinion and a personal opinion. When an expert has formulated an opinion, it is reasonably presumed that the expert has drawn upon information accumulated and published over the years. *The defining attributes of an expert opinion relate not to the credentials held by the individual whose fingers type the words or from whose mouth the words flow; rather, the requisite characteristics relate to the procedures that were employed in formulating the opinion and the body of knowledge that forms the foundation upon which those procedures were developed* [italics added]. If the accumulated knowledge of the expert's field was not utilized, the opinion expressed is not an expert opinion. It is a personal opinion, albeit one being expressed by an expert (Martindale, 2001).

Thus, a witness might well be qualified quite eminently as a clinical or forensic psychologist yet the opinion nonetheless be inadmissible because it rests in some part on a principle or technique that is not scientifically reliable and valid, such as concluding that a child has been sexually abused on the basis of play with anatomically detailed dolls (Lilienfeld et al., 2003, pp. 57–61), or upon unproven theoretical assumptions or personal values (e.g., preschool children need a primary home, siblings should not be separated, etc.). The question is not whether this person is qualified as an expert because of his or her credentials as a psychologist or psychiatrist but, rather, whether *any* mental health witness, however well credentialed, is competent to answer the particular question, for example, what arrangement is in the best interest of the subject child?

Otherwise stated, does this witness have "specialized knowledge" that allows each required inference to be drawn so that it may be stated to a "reasonable degree of *professional certainty*"? That question, in due course, turns on the underlying question as to whether the witness's profession has produced a reliable and valid knowledge base upon

which to predicate each inference. This issue, ultimately, turns on whether the empirical research supports scientifically reliable and valid constructs to support the inferences. The process of witness qualification establishes only that this particular witness is qualified to speak for his or her profession, to state in a courtroom what is known *empirically* to that profession. What is “known” to the mental health profession is that which is established empirically, not what this particular witness may have idiosyncratically intuited or concluded on the basis of personal value judgments.

OPINION ON ULTIMATE ISSUES

Even though evidence doctrine no longer precludes opinions on ultimate issues as a general matter, the courts ought not to permit them in custody cases as a matter of evidentiary law and the mental health profession should not provide them as a matter of professional ethics.

There are two related reasons why such opinions ought not to be permitted:

- They exceed the boundaries of the empirical knowledge base of the mental health profession.
- They implicitly misrepresent the limits of that knowledge base.

As one noted scholar expresses the problem:

Frequently, evidence that lacks reasonable scientific backing is presented to the courts as though it had scientific merit. Although corrupt individuals occasionally may act as “hired guns,” testifying in favor of those who retain them for a price, a much more common and pernicious problem is the uncorrupt but inaccurate expert who presents pet theories and personal prejudices as scientific fact (Baerger, Galatzer-Levy, Gould, & Nye, 2002).

The Melton treatise also makes the point quite cogently:

Despite the fact that such opinions are commonly sought and, unfortunately, are commonly given, *mental health professionals ordinarily should refrain from giving opinions as to ultimate legal issues*. . . . When experts give such opinions, they usurp the role of the fact finder by suggesting that the opinions are based on specialized knowledge specific to the profession.

Note in this regard that although Rule 704(a) allows experts to give opinions on ultimate issues, Rule 702 prohibits admission of any opinion not based on specialized knowledge, a prohibition which presumably can include ultimate issue opinions. Indeed, Rule 704(b) (an amendment to the original Rule 704 which was inspired by John Hinckley’s acquittal on insanity grounds) makes this point concretely with respect to mental state testimony in criminal cases. The position we take is that the same evidentiary prohibition should apply to *all* types of cases. In any event, even if a court permits such an opinion to be admitted as a matter of law, it should not be offered as a matter of professional ethics because of the implicit misrepresentation of the limits of expertise involved when a clinician acting as an expert witness gives a legal opinion in the guise of mental health knowledge (Melton et al., 1997, p. 17).

Lest there be any doubt that Melton et al. include mental health testimony in connection with custody opinions, they state quite compellingly:

Indeed, there is probably no forensic question on which overreaching by mental health professionals has been so common and so egregious. Besides lacking scientific validity, such opinions have often been based on clinical data that are, on their face, irrelevant to the legal questions in dispute (p. 484).

Indeed, whatever position one might take on the ultimate issue rule with respect to other species of expert testimony, such opinions by mental health witnesses on the ultimate question of a child's best interest ought not to be allowed. It is less an ultimate issue matter than it is a question of basis.

The best interests standard is a legal and socio-moral construct, not a psychological construct. There is no empirically supportable method or principle by which an evaluator can come to a conclusion with respect to best interests entirely by resort to the knowledge base of the mental health profession. In this regard, Melton et al. (1997) have noted that "there has been remarkably little research meeting minimal standards of methodological rigor about the effects of various *custody arrangements* on children and families of different characteristics" (p. 484). Krauss and Sales (1999) observed that ". . . few studies have focused on the directly relevant legal questions or used the methodological rigor that is necessary for a psychologist to offer exclusively science-based opinions on the best interest of a particular child" (p. 89).

O'Donohue and Bradley (1999) called for a moratorium on all participation by mental health professionals in the child custody arena because ". . . the mental health professional currently cannot ethically conduct these evaluations." In so doing, they refer to "the lack of empirical research that clearly identifies relevant variables involved in child custody and their possible causal or correlative relationships to key outcomes."

The custody court is confronted with the challenge of determining with which parent the child's interest will best be served. Sometimes this involves selection of the "better" parent. Too often, the court must choose the one who is less "bad." There is no empirically verified psychological construct of "good parent" or "bad parent", let alone a construct for the comparative analysis that the court is called upon to perform. While we have established certain correlations between parenting style and child outcome, "good" and "bad" are socio-moral constructs that, by definition, are not within the purview of psychology as a behavioral science. There is also a noted "lack of scientific methodology, empirical grounding, and psycho-legal relevance" in the custody evaluation field (Baerger et al., 2002, p. 36). Yet, these opinions are highly influential:

The recommendations contained in child custody evaluations ("CCEs") exert considerable influence on the course of ongoing custody litigation. Many courts accord significant weight to the opinions of child custody evaluators, often accepting the evaluator's recommendations without challenge. An evaluator's recommendations can also precipitate case settlement or material concessions once both parties become aware of the evaluator's findings (Baerger et al., 2002; Horvath, Logan, & Walker, 2002).

In effect, the legal, socio-moral construct of "best interest" has no "specific, operational definition" in terms of psychological assessment and is "so ambiguous and value laden that reaching a consensus that would permit accurate measurement of this characteristic in a forensic assessment" is not possible (Shuman, 2002, p. 144).

As one prominent writer in the field states:

We often use scientifically crafted data-gathering tools, such as standardized tests. We organize our data and form inferences based on theoretically developed models of human behavior that guide our conclusions. We are, in most aspects of our work, scientists struggling to fit our observations into models of behavior that can inform the court about what is in the child's best psychological interests.

We should not delude ourselves, however. Our techniques and models are only in their infancy. Our tests are often not validated on forensic populations, let alone the specific

populations of those undergoing custody evaluations. Our interview protocols often are home-grown, with little, if any, reliability data available concerning test-retest or internal consistency. Our recommendations are based not on empirically driven conclusions but on best-guess strategies.

For example, no empirical work has been done in which a matched set of children with similar test and interview data is placed randomly into different custodial arrangements to examine overall effectiveness of one placement over another. Such an experiment would be blatantly unethical and inhumane. We cannot manipulate families to serve our science. Yet we often provide testimony to the court about custodial arrangements with the arrogance of “true” science implied. We are but sophisticated guides for the trier of fact through a confusing array of psychological technology, a technology never intended for use in custodial assessment. Through our learned and judicious use of psychological theories, methods, and data gathering, we determine our best guesses possible. Our tools are often not valid for custodial assessment. Our models are often rationally, not empirically, derived. And our opinions are more educated guesses than truth. We need to be careful in how we present our data and opinions to the court so as not to mislead (Gould, 1998).

Psychology has a long way to go before it has empirically tested competing theories about the long-term adaptation of children under various access plans and, as noted above, the multidimensional complexity of such a task, along with our inability to randomly assign children to different custody conditions, make the likelihood of an adequate test extremely remote. Until custody-specific theory and replicated empirical research on this topic is a fact, rather than a hope, the custody evaluator has little empirically based guidance about what relevant variables to analyze, let alone their priority or relative weight. A uniform methodology for conducting custody evaluations likewise must await the empirical research that indicates what the relevant variables are that must be studied and development of a uniform method of assessing them. Absent that research, the evaluator cannot possibly draw the necessary inferences without seeking guidance from clinical experience, which is often nothing more than accumulated bias, personal moral/value judgments, or unverified theoretical assumptions (Wittmann, 1985).

The issues presented here go far beyond questions of “best practice” because one cannot begin to consider practice standards in the absence of an empirical knowledge base that informs the discussion. They are not problems that can be subsumed within the rubric of “it goes to weight” either, or the time-honored though doctrinally unsupportable, “I’ll take it for what it’s worth.” Rather, these issues go to the heart of the law of evidence, that is, relevance and reliability, and thus present a question not of weight, but of admissibility.

DISCUSSION AND CONCLUSIONS

Notwithstanding the strong case that exists that it is unethical for mental health witnesses to tender specific custody recommendations and that their admission at trial violates fundamental evidentiary doctrine, it would be a mistake to blame the current situation solely on the mental health profession. That profession is simply responding to the demand of the legal consumers, the lawyers and judges who actively seek such invalid opinions. Privately, many evaluators will acknowledge that they deliver these recommendations with great internal reluctance, but yield to consumer pressure for fear of being out of business if they refuse.

Psychologists frequently assume roles in which they’re expected to offer opinions, even though uncertainty is enormous and the scientific knowledge base is extremely limited. For example,

in custody evaluations that involve reasonably normal individuals, it's often impossible to know which parent will provide a child with a better environment. Research in this area is limited, and there are few or no well-validated procedures for arriving at recommendations.

. . . Nevertheless, the lawyers and judges in such cases often press psychologists for a firm opinion.

Under such demands, psychologists sometimes resort to unvalidated procedures. The reasoning seems to be that any test, even one of dubious validity, will perform better than merely flipping a coin (Wood et al., 2003, p. 297).

The legal system has both the right and the duty to exclude opinions that are not supported by good science. It is understandable that lawyers and judges actively seek recommendations that are not so supported. This reflects both the natural human desire for easy answers to difficult problems and the legal profession's lack of understanding of scientific method in general and, more specifically, its proper application in the behavior science milieu.

Lawyers and judges perceive a value to specific recommendations in large part because they often facilitate settlement without trial. It is important to understand, however, that the leverage the recommendation exerts on litigants has little, if anything, to do with its underlying validity. As a general proposition, litigants settle cases when they perceive that they likely will do worse if they go to trial. Thus, the typical scenario in custody litigation is that the evaluator's report comes in with a specific recommendation. The attorney for the disfavored party tells the client "the judge is highly likely to go with the recommendation." Loosely translated, "you are going to lose, so why not settle now and save yourself the risk and cost of trial." It can fairly be said that the client does not hear the recommendation and experience an epiphany, suddenly realizing that the other parent is indeed the better choice. But for the *perception* that the judge will credit the recommendation and act upon it, the report would exert little leverage on the settlement process. Hypothetically, the same leverage could be obtained if everyone believed that the judge had tossed a coin and the result would be determinative or highly persuasive. The distinction is that no one would believe the coin toss is a relevant predictor of parenting capability but many within the legal system, being uninformed of behavioral science precepts, do believe that the mental health witness's recommendation is.

Yet, specific recommendations can also have unintended effects. When a judge receives a specific custody recommendation from an evaluator and then decides that it ought *not* be followed, the recommendation may well pose an obstacle that the judge needs to surmount in writing the decision. Unfortunately, some appellate courts have given quasi-presumptive status to the opinions of court-appointed evaluators, mistakenly confusing court-appointed status with freedom from bias and other assessment errors. (*Rentschler v. Rentschler*, 1995; *In re Custody of Rebecca B.*, 1994, *but see, Chait v. Chait*, 1995; *Edgerly v. Moore*, 1996). Such decisions leave cautious trial judges feeling the burden of justifying any departure from the mental health opinion by "writing their way out of it."

Given the pervasiveness and potentially enormous impact of behavioral science testimony, it is imperative that lawyers and judges become more educated and informed consumers of mental health opinions. Advanced education in scientific method and empirical behavioral science research will position lawyers and judges to bring more qualitative analysis to bear upon forensic testimony and produce an essential understanding that the legal system ought not make demands upon behavioral science that it is presently unequipped to meet nor admit recommendations that are not predicated upon an empirically established specialized knowledge base.

RECOMMENDATIONS

As implied above, it is our position that clinicians with adequate forensic training can provide matrimonial and family court judges with very useful and helpful information at Levels I and II. For example, admissions by litigants regarding disputed questions before the court, compilations of descriptions regarding parenting behavior over time, and data based inferences regarding parenting skills/deficits. In addition, the articulation of a child's expressed concerns, emotional pain, developmental needs, and preferences all represent ethical contributions to the fact finding process as long as each conclusion that is drawn is based in a cautious hypothesis-testing model, convergent data, and the specialized knowledge of the profession. The simple listing for the court of psychological risk factors associated with various access plans (Level III) can also be done in an ethical manner as long as it is carefully grounded in case specific data and the specialized knowledge base of the mental health profession. In these areas, we believe that forensic reports can have very important value to the fact-finding effort and that they help to humanize an often excessively polarized and distorting adversarial process. However, our literature review suggests the following changes in professional practice:

- (1) Given our conclusion that psychologists can only offer ethical and helpful information at Levels I and II, and a small bit of useful information at Level III, the courts should continue to engage qualified and circumspect forensic evaluators to assist the fact-finder in bringing forth custody-relevant information. However, we do suggest a moratorium on the practice of psychologists providing recommendations regarding specific custody arrangements and schedules.
- (2) Psychologists should routinely include warnings in forensic reports about the very significant limitations in our capacity to predict the best or worst custody plans for a particular family as a step toward educating the judiciary (particularly if they choose to continue the practice of making specific custodial recommendations).
- (3) The American Psychological Association should work toward articulating a more clear position against specific custody recommendations (given that it is almost impossible to argue that adequate data exists to make such prescriptions).
- (4) Judges should begin to help the psychology discipline rein itself in by including in court orders a specific statement precluding the expert witness from addressing the ultimate custody-plan issue.
- (5) Clinicians should avoid addressing custody-related issues that are largely non-psychological (economic benefits of different child placements, qualitative differences between school districts, etc.).
- (6) Judges and lawyers should routinely be provided with more comprehensive training regarding seminal psychological and developmental issues relevant to access planning. More broadly, law school curricula ought to include basic and advanced courses to educate the legal profession with respect to the many areas where law and the behavioral sciences intersect.
- (7) Clinicians should be allowed to function as expert evaluators in custody matters only if they have had adequate forensic training (to help ensure respect on the part of the clinician for the critical importance of an empirical grounding for inferences when such inferences can result in the limitations on human liberties).
- (8) Future research should focus on developing valid and reliable instruments for measuring Level I and Level II variables. In addition, attempts should be made to study the multivariate combinations of individual, family, and access plan variables that tend to be correlated with negative and positive child outcomes (admittedly, a daunting task).

NOTES

1. See, for example, *Mtr. of Bennett v. Jeffreys*, 40 NY2d 543 (1976): “In custody matters parties and courts may be very dependent on the auxiliary services of psychiatrists, psychologists, and trained social workers. This is good. But it may be an evil when the dependence is too obsequious or routine or the experts too casual. Particularly important is this caution where one or both parties may not have the means to retain their own experts and where publicly compensated experts or experts compensated by only one side have uncurbed leave to express opinions which may be subjective or are not narrowly controlled by the underlying facts.”; and *Mtr. of Sayeh R. v. Monroe County DSS*, 91 N.Y.2d 306, 693 N.E.2d 724: “It is necessary also to make some reference to the role of the psychological material presented in the instant case. Experts, who predict future consequences based on their professional theories and examinations of subject children, should not be elevated to the singular importance of, in effect, overriding the array of pertinently balanced jurisdictional protections afforded to decrees affecting one of society’s most sacrosanct relationships—parent and child. Courts must beware lest the unique juridical authority to decide these cases be sacrificed to the sheer crosswinds of paid or even so-called independent experts.”

2. Even in the basic statement, the concept of “relevance” ought to keep out most of what passes in the trade today because proffered testimony cannot be relevant if it is not reliable. The two concepts are inextricably intertwined.

3. The issue of keeping unreliable evidence from prejudicing the factfinder is sometimes assumed to be less significant in custody cases because they are usually tried to a judge, rather than a jury. This notion is not well-founded:

Because judges, not jurors, are responsible for the final decision in these cases, the judges may believe that they will not be influenced by expert testimony that might fail the Daubert admissibility standard and, therefore, do not feel it is necessary to apply the Daubert standard to expert testimony prior to its admission into evidence in this context. Obviously, assumptions that judges can mentally differentiate good scientific expert testimony from bad scientific expert testimony, and differentially evaluate scientific and clinical testimony, are unanswered empirical questions. Likewise, the assumption that judges are less influenced by bad scientific testimony than jurors is also an unanswered empirical question. Yet, similar to the prior assumptions, it appears from the existing psychological literature on decision making that these assumptions are questionable at best (Krauss & Sales, 1999, p. 84).

4. The same text also notes that tensions between professional psychology and research psychology reached “crisis levels” by the late 1980s when practitioners took control of the APA, prompting research psychologists to break away in 1988 and form the American Psychological Society whose defined mission is “to promote, protect, and advance the interests of scientifically oriented psychology in research, application, teaching, and the improvement of human welfare.”

REFERENCES

- Ackerman, M., & Schoendorf, K. (1992). *ASPECT: Ackerman-Schoendorf Scale for Parent Evaluation of Custody—Manual*. Los Angeles: Western Psychological Services.
- American Psychological Association. (2003). Ethical principles of psychologists and code of conduct. *American Psychologist*, 57(12), 1060–1073.
- American Psychological Association. (1994). Guidelines for child custody evaluations for in divorce proceedings. *American Psychologist*, 49, 677–682.
- Baerger, D. R., Galatzer-Levy, R., Gould, J. W., & Nye, S. G. (2002). A methodology for reviewing the reliability and relevance of child. *Journal of the American Academy of Matrimonial Lawyers*, 18, 35–74.
- Bagby, M. R., Nicholson, R. A., Buis, T., Radovanovic, H., & Fidler, B. (1999). Custody evaluations detecting malingering and defensive responding on the MMPI-2 in family custody and access evaluations. *Journal of Personality Assessment*, 62(2), 191–203.
- Bauserman, R. (2002). Child adjustment in joint-custody vs. sole custody arrangements: A meta-analytic review. *Journal of Family Psychology*, 16(1), 91–102.
- Bow, J. N., & Quinnell, F. A. (2001). Psychologists current practices and procedures in child custody evaluations five years after American Psychological Association guidelines. *Professional Psychology: Research and Practice*, 32(3), 261–268.
- Bow, J. N., & Quinnell, F. A. (2004). Critique of child custody evaluations by the legal profession. *Family Court Review*, 42, 115–126.

- Brennan, J. F. (2003). *History and systems of psychology* (6th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Bricklin, B. (1989). *Perception of Relationships Test manual*. Furlong, PA: Village Publishing.
- Bricklin, B. (1984, 1990a). *Bricklin Perceptual Scales manual*. Furlong, PA: Village Publishing.
- Bricklin, B. (1990b). *Parent Awareness Skills Survey manual*. Furlong, PA: Village Publishing.
- Bricklin, B. (1992). Data-based tests in child custody evaluations. *American Journal of Family Therapy*, 20, 254–265.
- Brodzinsky, D. M. (1993). On the use and misuse of psychological testing in child custody evaluations. *Professional Psychology: Research and Practice*, 24(2), 213–219.
- Butcher, J. N. (1997). Frequency of MMPI-2 scores in forensic evaluations. *MMPI-2 News and Profiles*, 8(1), 2–4.
- Cummings, M. E., & Davies, P. (1994). *Children and marital conflict: The impact of family dispute and resolution*. New York: Guilford Press.
- Felner, R. D., Rowlison, R. T., Farber, S. S., Primavera, J., & Bishop, T. A. (1987). Child custody resolution: A study of social science involvement and impact. *Professional Psychology: Research and Practice*, 18, 468–474.
- Foster, K. R., & Huber, P. W. (1997). *Judging science: Scientific knowledge and the federal courts*. Cambridge, MA: MIT Press.
- Galatzer-Levy, R. M., & Kraus, L. (1999). *The scientific basis of child custody decisions*. New York: Wiley.
- Garb, H. N. (1998). *Studying the clinician*. Washington, DC: American Psychological Association.
- Goldstein, J., Freud, A., & Solnit, A. (1973). *Beyond the best interests of the child*. New York: Free Press.
- Gould, Jonathan W. (1998). *Conducting scientifically crafted child custody evaluations*. Thousand Oaks, CA: Sage.
- Heilbrun, K. (2001). *Forensic mental health assessment*. New York: Kluwer Academic.
- Hetherington, M., & Kelly, J. (2002). *For better or for worse: Divorce reconsidered*. New York: W.W. Norton & Company.
- Hetherington, M., & Stanley-Hagen, M. (2002). Parenting in divorced and remarried families. In M. H. Borstein (Ed.), *Handbook of parenting* (Vol. 3). NJ: Erlbaum.
- Horvath, L. S., Logan, T. K., & Walker, R. (2002). Child custody cases: A content analysis of evaluations in practice. *Professional Psychology: Research and Practice*, 33, 557–562.
- Hynan, D. L. (2003). Parent-child observations in custody evaluations. *Family Court Review*, 41(2), 214–223.
- Karra, D., & Berry, K. (1985). Custody evaluations: A critical review. *Professional Psychology: Research and Practice*, 16(1), 76–85.
- Karras, D., & Berry, K. K. (1985). Custody evaluations: A critical analysis. *Professional Psychology: Research and Practice*, 16(1), 76–85.
- Kelly, J., & Lamb, M. (2000). Using child development research to make appropriate custody and access decisions for young children. *Family & Conciliation Courts Review*, 38(3), 297–311.
- Krauss, D. A., & Sales, B. D. (1999). The problem of “helpfulness” in applying *Daubert* to expert testimony: Child custody determinations in family law as an exemplar. *Psychology of Public Policy and Law*, 5, 78–99.
- Lamb, M. E., Orbach, Y., Sternberg, K. J., Hershkowitz, I., & Horowitz, D. (2000). Accuracy of investigators verbatim notes of their forensic interviews with alleged child abuse victims. *Law and Human Behavior*, 24, 699–708.
- Lilienfeld, S. O., Lynn, S. J., & Lohr, J. M. (2003). *Science and pseudoscience in clinical psychology*. New York: Guilford Press.
- Lyon, T. D., & Koehler, J. J. (1996). The relevance ratio: Evaluating the probative value of expert testimony in child sexual abuse cases. *Cornell Law Review*, 82, 43–78.
- Martindale, D. A. (2001). Cross-examining mental health experts in child custody litigation. *The Journal of Psychiatry & Law*, 29, 483–511.
- Mayes, L. C., & Truman, S. D. (2002). Substance abuse and parenting. In M. H. Borstein (Ed.), *Handbook of parenting* (Vol. 4) (pp. 329–360). Mahwah, NJ: Erlbaum.
- McCann, J. T., Flens, J. R., Campagna, V., Collman, P., Lazzaro, T., & Connor, E. (2001). The MCMI-III in child custody evaluations: A normative study. *Journal of Forensic Psychology Practice*, 1, 27–44.
- Melton, G. B., Weithorn, L. A., & Slobogin, C. (1985). *Community mental health centers and the courts: An evaluation of community-based forensic services*. Lincoln, NE: University of Nebraska Press.
- Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C. (1997). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (2nd ed.). New York: Guilford Press.
- Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., et al. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, 56(2), 128–165.
- O’Donohue, W., & Bradley, A. R. (1999). Conceptual and empirical issues in child custody evaluations. *Clinical Psychology: Science and Practice*, 6, 310–322.

- Otto, R. K., Edens, J. F., & Barcus, E. H. (2000). The use of psychological testing in child custody evaluations. *Family & Conciliation Courts Review*, 38(3), 312–340.
- Patterson, G. R., & Fisher, P. A. (2002). Recent developments in our understanding of parenting: Bidirectional effects, causal models, and the search for parsimony. In M. H. Borstein (Ed.), *Handbook of parenting* (Vol. 5; pp. 55–88). Mahwah, NJ: Erlbaum.
- Pope, K. S., Butcher, J. N., & Seelen, J. (2000). *MMPI, MMPI-2, & MMPI-A in court: A practical guide for expert witnesses and attorneys* (2nd ed.). Washington, DC: American Psychological Association.
- Reagan, R. T. (1999). Book review—judging science: Scientific knowledge and the federal courts, by Kenneth R. Foster & Peter W. Huber. *Oklahoma Law Review*, 52, 291–301.
- Rothbaum, F., Weisz, J., Pott, M., Miyake, K., & Morelli, G. (2000). Attachment and culture: Security in the United States and Japan. *American Psychologist*, 55(10), 1093–1104.
- Rubin, K. H., & Burgess, K. B. (2002). Parents of aggressive and withdrawn children. In M. H. Borstein (Ed.), *Handbook of parenting* (Vol. 1). Mahwah, NJ: Erlbaum.
- Shuman, D. W. (2002). The role of mental health experts in custody decisions: Science, psychological tests, and clinical judgment. *Family Law Quarterly*, 36, 135.
- Strong, J. (Ed.). (1999). *McCormick on evidence* (5th ed.) (Vols. 1–2). St. Paul, MN: West Group.
- Teti, D. M., & Candelaria, M. A. (2002). Parenting competence. In M. H. Borstein (Ed.), *Handbook of parenting* (Vol. 4; pp. 149–180). Mahwah, NJ: Erlbaum.
- Tippins, T. M. (2003). Custody evaluations—Expertise by default? *New York Law Journal*, 230, 3.
- Turk, D. C., & Salovey, P. (Eds.). (1988). *Reasoning, inference, & judgment in clinical psychology*. New York: Free Press.
- Tutin, J. (1993). The persistence of initial beliefs in clinical judgment. *Journal of Social and Clinical Psychology*, 12(3), 319–335.
- Waxler, Z. W., Duggal, S., & Gruber, R. (2002). Parenting and psychopathology. In M. H. Borstein (Ed.), *Handbook of parenting* (Vol. 4; pp. 295–328). Mahwah, NJ: Erlbaum.
- Wigmore, J. (1979). *Wigmore on evidence*. Boston: Little Brown.
- Wittmann, J. J. P. (1985). Child advocacy and the scientific model in family court: A theory for pre-trial self-assessment. *The Journal of Psychiatry & Law*, 13, 61–82.
- Wittmann, J. J. P. (2003). Should forensic psychologists make custody recommendations? *New York Family Law Monthly*, 4(11), 1–3.
- Wood, J. M., Nezworski, M. T., Lilienfeld, S. O., & Garb, H. N. (2003). *What's wrong with the rorschach? Science confronts the controversial inkblot test*. San Francisco: Jossey-Bass.

Table of Cases

- Bennett v. Jeffreys, 356 N.E.2d 277 (N.Y. 1976).
- Chait v. Chait, 638 N.Y.S.2d 426 (N.Y. App. Div. 1995).
- Edgerly v. Moore, 647 N.Y.S.2d 773 (N.Y. App. Div. 1996).
- Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579 (1993).
- Daubert v. Merrell Dow Pharm., Inc., 43 F.3d 1311 (9th Cir. 1995).
- Frye v. United States, 293 F. 1013 (D.C. Cir. 1923).
- In re Custody of Rebecca B., 611 N.Y.S.2d 831 (N.Y. App. Div. 1994).
- In re Sayeh R., 693 N.E.2d 724 (N.Y. 1997).
- Kumho Tire Co., Ltd. v. Carmichael, 526 U.S. 137 (1999).
- Jones v. United States, 548 A.2d 35 (D.C. 1988).
- People v. Scarola, 525 N.E.2d 728 (N.Y. 1988).
- People v. Wernick, 674 N.E.2d 322 (N.Y. 1996).
- People v. Wesley, 633 N.E.2d 451 (N.Y. 1994).
- Rentschler v. Rentschler, 611 N.Y.S.2d 523 (N.Y. App. Div. 1995).
- State v. Montalbo, 828 P.2d 1274 (Haw. 1992).
- Turpin v. Merrell Dow Pharm., Inc., 959 F.2d 1349 (6th Cir. 1992).
- Tyson v. Keane, 991 F.Supp. 314 (S.D.N.Y. 1998).

Timothy M. Tippins, Esq. is an Adjunct Professor of Law at Albany Law School and a private practitioner who has engaged in matrimonial and family law practice for over thirty years. He now devotes his practice time to serving as special counsel and consultant to other family law practitioners on a nationwide basis, serving as trial and appellate counsel, with special emphasis on the presentation and cross-examination of expert mental health testimony. Tippins has served in all major professional leadership positions in the New York family law community, including President of the American Academy of Matrimonial Lawyers—New York Chapter, Chair of the New York State Bar Association Family Law Section, and Chair of the New York State Bar Association Task Force on Family Law. Tippins is a regular feature columnist for the New York Law Journal, where his numerous articles regarding custody evaluations have been published. He has lectured before many major professional organizations, such as the Association of Family and Conciliation Courts, American Academy of Matrimonial Lawyers, ABA, NYSBA, and the New York State Psychological Association. Together with Dr. Wittmann, he has taught the Tippins-Wittmann Empirical Forensic Model at the NYS Judicial Institute and at other forums around the country.

Jeffrey P. Wittmann is a licensed psychologist, trial consultant, and divorce mediator whose practice concentrates on consultation in contested custody and access matters. He serves as a trial consultant for major law firms in the Manhattan area and elsewhere. Dr. Wittmann is a nationally recognized expert on issues related to the intersection of law and psychology and on professional practices in the child custody area, is regularly on the faculty for training seminars offered to attorneys and mental health professionals, and has been a frequently utilized scholar-trainer for judges in Family, Supreme, and Appellate courts. He lectures on forensic and divorce psychology at both the undergraduate and graduate level. Dr. Wittmann has provided invited testimony to the New York State Senate regarding matrimonial statutes; provides opinions on divorce-related matters to print, television, and radio media; and has served as a program consultant to ABC News. He is on the editorial board of the New York Family Law Monthly and is the author of Custody Chaos, Personal Peace (Perigee, 2001) and of numerous scholarly articles regarding forensic psychology.