

WHAT YOU SHOULD KNOW ABOUT THE FRYE STANDARD, THE FEDERAL RULES OF EVIDENCE, AND THE DAUBERT CRITERIA WHEN YOU GO TO COURT

THE ADMISSIBILITY OF EXPERT TESTIMONY, INCLUDING TESTIMONY ON SCIENCE, BY MENTAL HEALTH PROFESSIONALS

Barry Bricklin, Ph.D.

and

Gail Elliot, Ph.D.

Please Note: Any mental health professional who testifies in court will find helpful pointers in this paper. But it is also important that you, the mental health professional, request that your attorney read it. Many of the action steps required to make your testimony more effective will need to be initiated by your attorney.

Abstract: Following a discussion of the Frye Standard, the relevant Federal Rules of Evidence, the Daubert opinion and dicta, as well as expert testimony by mental health professionals, **this paper specifically addresses how *The Perception-of-Relationships Test (PORT)* and *Bricklin Perceptual Scales (BPS)* meet the criteria for admissibility of scientific evidence by demonstrating the following:** Both tests are testable and have been tested, with a combined database of 3,880 subjects. The tests have been published and peer-reviewed. They have a known error rate, and the predictive accuracy of the tests ranges between 88 and 92 percent agreement between test-based assertions and a wide range of validating criteria. A study found the tests to be the most widely used custody tests for children in the country, indicating that they have found general acceptance. Many test manual updates plus current revisions of the manuals have provided test users with information to maximize reliability and validity of the tests.

Village Publishing
73 Valley Drive
Furlong, PA 18925
800-553-7678 800-633-7223

COPYRIGHT 2009 THIS PAPER MAY BE FREELY COPIED OR TRANSMITTED BY ANY APPROPRIATE MEANS AVAILABLE.

WHAT YOU SHOULD KNOW ABOUT THE FRYE STANDARD, THE FEDERAL RULES OF EVIDENCE, AND THE DAUBERT CRITERIA WHEN YOU GO TO COURT

THE ADMISSIBILITY OF EXPERT TESTIMONY, INCLUDING TESTIMONY ON SCIENCE, BY MENTAL HEALTH PROFESSIONALS

**Barry Bricklin, Ph.D.
and
Gail Elliot, Ph.D.**

Please Note: Any mental health professional who testifies in court will find helpful pointers in this paper. But it is also important that you, the mental health professional, request that your attorney read it. Many of the action steps required to make your testimony more effective will need to be initiated by your attorney.

Abstract: Following a discussion of the Frye Standard, the relevant Federal Rules of Evidence, the Daubert opinion and dicta, as well as expert testimony by mental health professionals, this paper specifically addresses how *The Perception-of-Relationships Test (PORT)* and *Bricklin Perceptual Scales (BPS)* meet the criteria for admissibility of scientific evidence by demonstrating the following. Both tests are testable and have been tested, with a combined database of 3,880 subjects. The tests have been published and peer-reviewed. They have a known error rate, and the predictive accuracy of the tests ranges between 88 and 92 percent agreement between test-based assertions and a wide range of validating criteria. A study found the tests to be the most widely used custody tests for children in the country, indicating that they have found general acceptance. Many test manual updates plus current revisions of the manuals have provided test users with information to maximize reliability and validity of the tests.

THE FRYE STANDARD

In 1923, the D.C. Circuit Court was asked to address the issue of whether the polygraph procedure (the “lie detector”) met the standards of adequate science. The resulting opinion stated that for scientific evidence to be admissible in federal court, the method “must be sufficiently established to have gained general acceptance in the particular field in which it belongs” (Frye v. United States, 293 F. 1013, D.C. Cir. 1923, p. 1014. See also Groscup, J.L., Penrod, S.D., Studebaker, C.A., Huss, M.T., and O’Neil, K.M. The effects of Daubert on

the admissibility of expert testimony in State and Federal criminal cases. *Psychology, Public Policy and Law*, 2002, vol. 8, No. 4, pp. 339-372).

However, as years passed, scientific theories and newer technologies were developing at an expanding and quite rapid rate. The Frye Standard was seen as excluding valid and reliable procedures---procedures that just had not been around long enough to find “general acceptance” (and no court had ever operationally defined what the phrase “general acceptance” actually involved). Note that some of Einstein’s theories took more than 30 years to validate. Often, the kind of tools needed for such validation were simply not yet invented. George Gilles de la Tourette identified the syndrome named for him in the 1890’s, but it took more than 100 years for it to be accepted as a meaningful syndrome by the healing professions. The doctor who posited the role of prions as the cause of serious neurodegenerative diseases was ridiculed for 30 years until it was generally agreed by the scientific community that he was right all along. The same time patterns were true concerning arguments over the very existence of radon, the HIV and Lyme’s Disease. Also note how many useful theories there were on small particle phenomena (the search for the ultimately “smallest” building blocks of the universe) that could not be validated until societies, dozens of years later, could afford and knew how to construct multi-billion dollar particle accelerators (Segre, G. *Faust in Copenhagen: A struggle for the soul of physics*, 2007, p. 217).

Another flaw in Frye was that there were many procedures “generally accepted” in different fields as valid, that turned out not to be (think historically of blood-letting and nowadays about the individual application of criminal profiles).

Hence, in 1993, the Supreme Court, in the Daubert opinion, tried to address these, and similar, concerns (Daubert v. Merrill Dow Pharmaceuticals, Inc. 509 U.S. 579, 113 S.Ct. 2795, 1993).

DAUBERT AND BEYOND

Many of the opinions expressed in Daubert are straightforward and compelling (i.e., set clear precedents). Others (e.g., the dicta) are considered persuasive but “lack the force of an adjudication” (*Black’s Law Dictionary*, p. 454).

The straightforward findings (or “holdings”) are these.

1. The *Federal Rules of Evidence* (FRE), not Frye, provide the standards for admitting evidence in a Federal Court. (State systems have the option to adopt, not adopt, or to selectively adopt, certain federal rulings.) “General acceptance” is no longer to be a necessary precondition for admissibility.
2. The FRE (See the Federal Rules of Evidence. 1975. Pub.L. No. 93-595. Stat 1926; also see 28 U.S.C.A. 2072-2074) require that evidence be relevant

to the issues at hand, defined as whether or not the evidence has “any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence” (Rule 401, FRE). (See Mark, M.M. Social science evidence in the courtroom: Daubert and beyond. *Psychology, Public Policy and Law*, 1999, Vol. 5, No. 1, 175-193.)

3. Daubert clearly affirms the role of trial court judges as “gatekeepers” of the effort to limit scientific evidence to evidence that stems from reliable scientific procedures. The Daubert court recognized that many judges are not trained in science. In an effort to assist trial court judges, guidelines, called “dicta,” were expressed. Dicta “have persuasive power in making an argument, but they are not binding as precedent” (*Dictionary of Law*, 1996, Merriam-Webster, p. 1380. See also *Black’s Law Dictionary, 6th Edition*, 1990, West Publishing Co., p. 454 and Gifis, S.G. 2003, *Law Dictionary*, Barron’s Publishing, p. 146.) Later, we will present commentary by some of the nation’s foremost attorneys who are also doctoral level scientists on some key implications of these dicta.

The purpose behind the dicta was to help courtroom judges adhere to the most challenging holding of the Daubert court, the need for scientific evidence to be “reliable.” By this, the Supreme Court is referring to “evidentiary reliability---that is, trustworthiness... In a case involving scientific evidence, evidentiary reliability will be based upon scientific validity” (Mark, M. see citation above, p. 177). The term “reliable” as used in the legal system, translates to the world of mental health professionals as referring to both reliability and validity. The issues of relevance and probative value (the relative weight or value assigned to a piece of evidence in proving something to be, or not to be, true) are separate matters.

THE DAUBERT DICTA

Depending on how the ideas in the dicta are grouped, they generally appear in the literature as a list of either 4 or 5 or 6 items. We will group them in 5 categories.

1. Can the theory and/or technique be tested and/or falsified, and have they been? Please note carefully that the notion of falsifiability, based on ideas by Karl Popper (*The Logic of Scientific Discovery*, 1959) is generally misunderstood by everyone---judges, attorneys and most mental health professionals. It does not mean, as the notion is usually expressed, that some researcher has conducted a study that attempted to prove the methodology in question to be “false.” Almost every research endeavor does that. Thousands of concepts and tools have been found to be “false” in some study or another. And yet countless other scientists reject these research findings of “false,” asserting instead that the *purportedly “negative” research was false*, e.g., investigated the wrong concept or a poorly defined one, or employed tools and

methods that could not adequately or appropriately measure the scientific issue at question, or used inappropriate empirical equivalents. The latter are the manifestations in the real physical world that represent concepts. Einstein famously reiterated that empirical equivalents, that which is selected to represent a concept, are “free creations of the human mind” and that “physical facts are not, however it may seem...determined (in whole) by the ‘external world’” (Einstein, A. *Physics and reality. Journal of the Franklin Institute*, 1936, 221(3), pp. 1-34; Einstein, A. *Out of my later years*. 1984, pp. 98, 27, 274.) Also see Bricklin, B. & Bricklin, P.M. Custody data as decision-theory information. *Clinical Psychology: Science and Practice*, 1999, 6(3), 339-343.

Our main point is that when it is claimed that a method is either not testable or not falsifiable, the meaning is that the method *is inherently* not testable or falsifiable, either because the concepts or their empirical equivalents are so vaguely spelled out that it is *impossible to design a meaningful research study*. The classic example is the theory that whatever happens is God’s will. This is an untestable and non-falsifiable theory. Psychoanalytic analyses have been similarly faulted, since in many instances the causes of what occurs are stated *after the fact* of an occurrence. Psychoanalysts often feel free to explain whatever happens by assembling some list of pre-conditions that they claim must have been at work to result in whatever occurred. It is also worthy of note that Thomas Kuhn, in his seminal 1962 book, *The Structure of Scientific Revolutions*, (Chicago: University of Chicago Press) reminds us that the huge majority of scientists have totally ignored Popper’s paradigm, which actually deals more with ultimate aspirational “truth,” than with usefulness in the real world.

Summarizing, this guideline asks if the method in question *is capable* of being tested, and whether or not this has been done.

2. Guideline 2 asks whether the method has been published, and whether it has been peer-reviewed.
3. Another guideline asks if the method yields a known error rate.
4. One Daubert dictum retains the (still vaguely articulated) Frye standard of “general acceptance.”
5. Guideline 5, which does not often appear in the writings of commentators on Daubert, asks whether the method spells out techniques, advice or cautions that can be applied when the method is used, so that its reliability and validity is maximized.

IMPORTANT ISSUES REGARDING THE ADMISSIBILITY OF SCIENTIFIC EVIDENCE

Daubert was modified by the Joiner Decision (General Electric Co. v. Joiner, 522 U.S. 136, 1997) and the Kumho decision (Kumho Tire Co., Ltd. v. Carmichael, 119 S.Ct. 1167, 1999). The former claims that unless “abuse of discretion” occurs, appellate courts should not second-guess the admissibility decisions of trial courts, and that trial courts do not necessarily have to employ the Daubert dicta in their decision-making. Kumho attempts to shed light on the use of Daubert criteria in regard to expertise that may stem from clinical and/or other experience, and not necessarily from scientific theory or scientific method. It urges the use of Daubert criteria in areas where many psycholegal commentators claim their application is unclear.

The nation’s foremost psycholegal authority, Professor Bruce D. Sales, educated and trained as both an attorney and doctoral level psychologist, along with Daniel W. Shuman, another noted expert, offer the following points on Daubert, Joiner and Kumho (Sales, B.D. and Shuman, D.W. The impact of Daubert and its progeny on the admissibility of behavioral and social science evidence. *Psychology, Public Policy and Law*, 1999, Vol. 5, No. 1, 3-15.)

1. Kumho confounds rather than clarifies admissibility criteria by encouraging courts to use the Daubert criteria, developed to assess the tools of scientists, in cases where the evidence did not necessarily arise from the use of scientific methods e.g., clinical testimony. Clinical testimony is not always or even usually based on published data bases and clinical testimony is hardly ever in its totality based on known error rates.

2. The authors ask what “peer review,” (a flawed and largely political process to begin with) can possibly mean in regard to someone’s clinical opinions. When someone offers a clinical opinion in court, that opinion could not possibly have been peer-reviewed. Even negative *scientific* reviews, are often the result of a “small but determined cadre of critics...” who manage to become the gatekeepers of scientific journals (Weiner, I.B. 2001, Assessment advocacy, *Social Personality Assessment Exchange*, Vol. 12, No. 1, p. 7). A typical case of the potential bias inherent in peer review is that of physicist Lise Meitner. Strolling with her nephew in a wooded part of Sweden on December 24, 1938, she had the sudden insight that “when struck by a single neutron, a uranium nucleus can split into two pieces, releasing a considerable amount of energy (cite below).” The name “fission” was invented to explicate this process and, of course, it proved to be the key concept in developing the atom bomb. Niels Bohr, the universally revered physicist, argued that Lise Meitner should be awarded not only scientific recognition for her insight, but even the Nobel Prize. (She was not.) Her treatment at the hands of a large segment of the scientific community is considered a typical example of “wrongheaded political interference” that all too frequently can be inherent in peer review. Another example concerns Enrico Fermi, who was awarded the Nobel Prize in physics in 1938. Considered a “double-threat” scientist by his peers for skills as both an elegant theorist as well as an exemplary experimentalist, he wrote an article in

1933 on “beta decay.” His peer reviewers (for the journal, *Nature*) refused to publish it, claiming it contained “conjectures” far “removed from...reality.” It is now considered the “cornerstone” article in the field of elementary particles. (Segre, G. (2007). *Faust in Copenhagen: A struggle for the soul of physics*, pp. 36, 235, 236).

Other commentators make the following points.

1. Daubert offers no guidance on how the dicta should be aggregated (weighted, prioritized).
2. Daubert does not even require that a trial court must use *any* of the dicta. Mark (citation above, p. 178) says: “Use of these factors is not required.” Citing the decision, he quotes as follows: “...the trial judge must have considerable leeway in deciding” if evidence is reliable. J.P. Lipton (The use and acceptance of social science evidence in business litigation after Daubert. *Psychology, Public Policy and Law*, 1999, Vol. 5, No. 1, p. 65) says: In *Kumho*, the U.S. supreme Court reaffirmed that the Daubert factors “do not constitute a definitive checklist or test” and were meant to be “helpful not definitive.”
3. Lipton (cite above, p. 64) further notes that, in *Daubert*, the Court wrote that the fact of publication (or lack thereof) in a peer-reviewed journal is a “relevant, though not dispositive” consideration in deciding scientific validity.

DAUBERT IN MORE DETAIL

One of the main findings in *Daubert* was that the Federal Rules of Evidence (FRE), supersede the *Frye* criterion. Therefore, it is helpful for any mental health professional who is called upon to testify as an expert in court to be aware of key FRE items.

Many legal authors believe FRE 402 is a “starting point” for a *Daubert* analysis. Lipton (citation above, p. 61-62) says: “...all relevant evidence is admissible” and “evidence which is not relevant is inadmissible.” FRE 401 defines relevant evidence as “that which has any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.”

FRE 403 warns that the value to the courts of “opinion testimony” must be greater than its possible “prejudicial” impact on a decision maker. (See our text, *The Custody Evaluation Handbook*, 1995, Brunner/Mazel, pp. 5-11 for a more complete discussion.)

FRE 702, 703, and 704 are of particular importance to mental health professionals, since they spell out the definitions of expertise, and hence co-

determine whether any given mental health professional will or will not be accepted by the court as an expert.

FRE 702 is so important that we will return to it later. FRE 702 states the basic rule on the admissibility of expert testimony: “(I)f scientific, technical, or other specific knowledge will assist the trier of fact (the judge or jury) 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.” In order that testimony from an expert be admitted, it is necessary that the judge or jury should need assistance concerning the issues at hand, the scientific foundations permitting the testimony should be sufficiently reliable to permit generalizations to be made on their bases, and the specific witness must be qualified as an appropriate expert on the subject. (We will return to the issue of the reliability of evidence later for it is here we deal with the more controversial nature of the dicta.)

FRE 703 states an expert’s opinion can be based on any of the following: direct observation; facts acquired by hearing (or reading) the testimony offered at a given trial; materials that are obtained through books, literature, or “...experience to which the expert has access.” Further, an expert may rely on hearsay or evidence ordinarily not permissible in a courtroom, “if it is of the type ordinarily relied on by persons in his profession.” (Cites from Bricklin, 1995, see above.)

A key issue, on which participants in the legal process still do not agree, has to do with the area of the so-called “ultimate issue.” In a custody context, this would mean that while an evaluator who served as an expert witness could offer facts and information, and certain (limited) conclusions, about a given case, he or she could not address the “ultimate” legal issue, that is, who should serve as the primary caretaking or custodial parent. (An “ultimate issue” is essentially the legal matter that must finally be answered or resolved in a given case.)

FRE 704, engendered to deal with the “ultimate issue” matter, says that expert testimony that in fact is given in the form of an opinion or inference would be legitimate, even if it embraced an ultimate issue that was to be eventually decided by the trier of fact.

SOME EXCEPTIONALLY IMPORTANT POINTS ABOUT YOUR QUALIFICATIONS AS AN EXPERT

Go back and read FRE 702. Please note, and make sure the judge notes, that your expertise likely stems from every single category mentioned: scientific knowledge, technical knowledge and experience, very specific knowledge, plus your skill sets, your overall experience in countless categories related to the field of which you are a member, and your training and education. Of interest is

that in a study concerning analyses of expert testimony involving 693 cases as considered by appellate judges, four categories of expertise were noted: Medical/Mental Health; Technical; Scientific; and Business. "Scientists" accounted for only 19.6 percent of the total sample (Groscup et al., 2002, citation above, p. 346).

Those who practice in the healing professions---physicians, psychologists, clinical social workers, marriage and family counselors---rarely, if ever, could give meaningful and helpful testimony as though it derived from any one category, e.g., "science." When a physician orders a radiographic film, or a psychologist administers and interprets a psychological test, much more than "science" is needed to interpret the results *in any individual case*.

Anyone who has ever been both a clinical practitioner, who uses tests, as well as a researcher-author, who has researched and created tests, knows that it is almost impossible to put together published statistical data bases that can address, and protect against, every single situation that can have a negative impact on a test's accuracy. Any of the following conditions can cause bias in most psychological measuring instruments: anxious frustration, sadness, tiredness, poor reading skills or undiagnosed learning disabilities. Further, a respondent may have accidentally (or deliberately) had a chance to read over the items on a test ahead of time or may have taken the test in the past and not mentioned this to the tester. This could yield a so-called "practice effect" that could significantly enhance a score. All of this means that the methods of science can rarely, by themselves, be used to make the decisions on how scientific data can be safely used. Clinical experience is almost always needed.

Further, in the healing professions, as in others, a piece of information, scientifically derived or not, is rarely used as a lone silver bullet. And the most complex process in *any* applied science is that of aggregation, the process in which the decision maker must put together (weigh, prioritize) multiple pieces of evidence. There is hardly ever a clearly defined scientific model (or algorithm) to inform this task, for the process involves looking at all the data interactively and collectively, and it is rare for the "collections" to be exactly the same in all cases that on the surface may seem similar. Further, one piece of evidence can change the meaning of another. We consulted in a case where someone's responses to a test were all excellent. The evaluator gave him a perfect score. When we reviewed the answers, they seemed "too good." We inspected the manual for this test and realized the respondent had memorized the answers, since they were absolutely word-for-word from the manual. Our awareness was not science in action; it was clinical intuition, based on years of experience. The task of assigning credibility to evidence, trust in the accuracy of one's data and especially relevance to all the items looked at, can hardly ever be done according to a hard-and-fast scientific model, nor can one usually "add up the pieces" with some algorithm. This process requires a good bit of clinical judgment. Note well then that your expert testimony will hardly ever stem from

one category (e.g., “science”). You will, so to speak, be constantly “changing hats” as you testify, and the expertise needed for helpful testimony will likely involve skills sets and knowledge based on all five categories mentioned in FRE 702.

Following, is a good example of the artificiality of the separation among all of the categories mentioned in FRE 702, i.e., scientific, technical, specific knowledge, skills, experience, training and education. Guglielmo Marconi was the man who developed the wireless telegraph. He was not an accepted member of any scientific field. He had no recognized formal education in any field. Those who knew him best called him a “tinkerer” (Larson, E. *Thunderstruck*. 2006, p. 17). While he struggled for years to develop the wireless transmission of sound, the scientific community, especially physicists, labeled him a fraud and a liar. Ironically, to say the least, he was awarded the Nobel Prize *in physics*. At his Nobel lecture in Stockholm (December 11, 1909, Marconi Fondazione), he himself said he never considered himself a physicist, and that he did not understand the scientific principles behind his invention and further conceded he was “not even a scientist” (Larson, cite above, p. 312). If he had to testify in a court as an expert, he would probably have to be qualified on the basis of his experience, since he readily admitted he had no real understanding of the scientific theories that could be put forward to explain his achievements. He said that he knew nothing “scientific” about wavelength oscillations, nor how to measure them, nor even which ones his method used. (Of course, eventually many of these concepts were explained to him by persons who did understand them.)

THE DAUBERT DICTA AND PERCEPTION-OF-RELATIONSHIPS TEST (PORT) AND THE BRICKLIN PERCEPTUAL SCALES (BPS)

Daubert insists that expert testimony, whether scientific or clinical, be “reliable.” The dicta are meant to assist judges in their assessment of reliability. (Remember that the legal term “reliable” would translate into the psychological world as referring to both validity and reliability.)

Please note that “dicta” have “persuasive power” but are not precedent-setting. Further, neither Daubert, Joiner or Kumho spell out how many should be used in a given case, nor how they should be weighted in terms of their usefulness in evaluating scientific reliability, nor even if they need to be used at all. Indeed, appellate judges who were reviewing challenged testimony from 693 experts, rated the Daubert dicta as the “least important” criteria to apply in their reviews, and the FRE (especially 702) as “most important” (Groscup et al., p. 365).

Having said all this, please note that the PORT and BPS successfully meet all 5 of the dicta previously listed.

1. (a) Is the theory or method testable? The answer is yes, the methods are testable. Both the PORT and BPS are based on clearly defined concepts and empirical equivalents. Both are designed to be parts of overall comprehensive evaluations, since their data are related to only some aspects of “best interests” considerations, albeit important aspects. They were never intended to be used as stand-alone silver bullets. The PORT and BPS are to be used to generate clinical hypotheses that can be cross-checked and then integrated with all of the information gathered in any comprehensive evaluation of which they are part. Their manuals and updates suggest the comprehensive evaluation include interviews, observations, document reviews, as well as other psychological tests. The PORT and BPS are meant to assist the evaluator to better understand how a particular child assigns value to each parent in a variety of family systems configurations. “Value” is measured by the degree to which interactions between a parent and a given child lead to comfortable and competent behaviors on the part of the child. How these concepts were operationalized is spelled out later.

1. (b) Have the methods been tested: The answer is yes. Both tests have been repeatedly tested, from the 1960’s into the 1990’s, yielding a combined data base of 3,880 cases.

The best way to understand the degree to which a test can be helpful to a decision maker is to examine the entire chain of reasoning that links evidence to conclusions. In the field of psychology, this would mean the method has to declare a concept of psychological interest e.g., intelligence, psychopathology, or here, useful child custody information. A central conceptual measurement target of the PORT and BPS is that of the value (or utility) to a child of parental behaviors, i.e., not what parents know, say or do, but rather the *impact* these factors have on a given child. Next, the researcher would have to define or specify the empirical equivalents of the concept. These would be what one looks for in the world of human sensory experience---experience that is available to all rational and sensate humans---that represent the concept.

For the concept of intelligence the sensory data may consist of performance in clearly defined areas, or school grades, or range-of-achievement in multiple life areas and so forth. For “psychopathology,” the empirical equivalents could be symptom lists, the number of periods of perpetually felt stress and the like. In scientific models, it is helpful if these sensory manifestations can be quantified. With a test, one set of empirical equivalents are the items selected to be in the test. The other more critical ones are those things one looks for in the world of sensory experience that represent what the researcher seeks to co-predict, in the case of the PORT and BPS, carefully defined instances of comfortable and effective behaviors on the part of a child following interactions with a given parent. When the test scores achieve a proper degree of association with the real-world manifestations the tests are designed to elucidate, the test has been

scientifically validated. (See Bricklin and Halbert, 2004a, for a more complete discussion of how Einstein, Infeld, Bohm, Piotrowski and others have similarly described this process of linking evidence to conclusions.)

With such an analysis, the decision maker can decide if the method uses concepts and empirical equivalents (the physical manifestations of concepts in the real sensory world) that reflect how the decision maker himself or herself thinks about the area of interest, in this case, informed child custody assessment.

The aforementioned chain of reasoning is spelled out in greatest detail in two articles which appeared in the *American Journal of Family Therapy*, vol. 32, pp. 119-138 and pp. 189-203. (The authors are Barry Bricklin and Michael Halbert.) A concise but fairly complete version appears in a book chapter entitled "Empirically Assisted Assessment of Family Systems," by Barry Bricklin and Gail Elliot (in *Divorce and Custody: Forensic, Developmental, and Clinical Perspectives*, edited by Linda Gunsberg and Paul Hymowitz, The Analytic Press, 2005, pp. 201-219). The following excerpts are from that book chapter.

At the heart of the PORT and BPS is the core concept developed in the Uniform Marriage and Divorce Act, section 203, which directs that a decision-maker should evaluate parenting factors that directly and possibly indirectly have an impact on the child. It is clear that (except at extreme ends of a continuum) "parental competence," as a concept, cannot be applied to any single individual, that is, "parental competence" does not "reside" in a parent. *Parental competence can only be understood as the property of a specific system.* For example, an abrupt, perfunctory style on the part of a father may greatly bother one child and hardly be noticed by another. The former child likely assigns meanings such as "I guess Dad doesn't have much time for me" to such utterances, while the latter child does not. In fact, if the latter child has information-processing strategies that work well only with short rather than lengthy communications, the child would actually prefer such communications. To this child, brevity constitutes either a useful cost/benefit ratio for utterances seen as a bit short on positive affect, or, at some deeper level, this child may even symbolize such behavior as a deep respect for the child's ability to fill in the "gaps" on his or her own.

The *Perception-of-Relationships Test (PORT)* and *Bricklin Perceptual Scales (BPS)* aim to provide data-based assistance in understanding the impacts different caretakers have on a child in selected family systems (Bricklin, 1984, 1989, 1992, 1997, 1998, 1999, 2002; Bricklin and Elliot, 2001, 2002 a, b; Bricklin and Halbert, 2004a, b). Their theoretical bases derive from systems concepts. A system must have two or more elements and each element must have an effect on the *whole* system. The elements (and their effects) are not only interdependent but, however subgroups form, none can have an independent effect on the system as a whole. Within this definition of a system,

one cannot affect a system of which one is part and then not oneself be impacted by this effect (Ackoff, 1999, pp. 15-17). Most people are more used to mechanistic than systems models. As Ackoff reminds us, the former approaches understanding and/or prediction via a deconstruction process, in which the elements are analyzed one by one, after which their interactions with other elements are analyzed. Ackoff asserts that with systems, synthesis must come before analysis. This is why mechanistic models are more concerned with structures and systems analyses with functions, the former with “knowledge,” the latter with “explanation.” Systems complexities affect both validation issues (some predictions will be true only in limited contexts), and the choice of one’s measurement reference standard, as well as measurement units. The units must be adequate (i.e., not introduce bias or inaccuracy) for a test’s aims. A reference standard is the entity to which a measurement score is compared in order to derive relevance for a specific decision. The reference standard is a previously examined group (normative paradigm), a previously examined individual (the single-participant paradigm), or it may be criterion-referenced (arbitrary). To the extent that mental health professionals think of systems at all, they think (incorrectly) of models in which stable traits interact. This manner of thinking is evident in the way they write their evaluations, with sections called “Mr. Jones,” “Ms. Jones,” child “Mary Jones,” child “Sam Jones,” as though one can assess each element in a system as a separate entity and then somehow add up the parts. *In systems-based decisions, the elements of the system cannot be evaluated apart from the interactions of those elements within the system.* Further, as systems change, the relevant reference standard can shift. In a child custody context, the way Child 1 assigns value to his or her parents is not always meaningfully comparable to the way Child 2 or 3 would assign value either to his or her own parents or to the parents of Child 1. Except perhaps at extremes, knowing how a parent compares to other parents tells little about the unique and specific value a parent has to a particular child. To measure an individual’s unique assignment of value, a single-participant reference would be used. However, in other circumstances, including “termination of parental rights” cases, it might be useful to use a group reference standard (although there is currently no accepted “minimal parenting” standard). Further, if a decision is to be made based on the extent of the discrepancy between how a given child assigns value to one parent as opposed to the other across a number of life areas and how children in general assign differential value to each parent, a group reference standard could be relevant. Ordinal data can adequately address the issue of how a child assigns value to his or her caretakers in specific life areas, while interval measures are needed to address how a child assigns differential value to his or her parents across multiple life areas. The decision to use an ordinal scale ($A > B$; $B > A$; $A = B$) for the BPS Point Score Scale was deliberate. For one thing, it would be difficult to create a suitable measurement unit---except grossly---by means of which a child could express differential value between A and B in circumscribed areas, especially if the values for each are similar. *More important, the use of an ordinal scale was meant to reduce the*

influence on a decision-maker of parental value to a child in any single life area. The ordinal measures are subsequently summed, so that the parent-of-choice suggested by a test is based on an interval scale that reflects a child's assignment of value across multiple life areas, seven with the PORT and 32 with the BPS. In other words, we wanted the tests to help identify the parent who had greater value to a child in many life areas, rather than who was "good" or "bad" in some few areas (although for clinical purposes only, BPS Point Scores will reflect the latter).

Note that systems complexities can have profound effects on the choice of validating criteria. The parent from whom a child seeks emotional closeness and/or active help can change dramatically depending on the family subsystems in which the child-parent interactions take place (Bricklin and Elliot, 2002a, b; Bricklin and Halbert, 2004b).

Our data-based tests assess the degree to which child-parent interactions lead to emotionally comfortable and behaviorally competent behavior on the part of the child in various family systems. The PORT generates data relevant to several systems, while the BPS is relevant to dyadic systems. They also seek ways to understand useful cost/benefit ratios. An anxious child may need that parent who can best calm him or her down, even though this parent may be less good at modeling competency, while such a pairing may be a relatively poor one for an already secure child.

The original and much of the subsequent validity research with the PORT compared PORT estimations of the particular parent with whom a child more comfortably and efficiently shared information than with the other to estimations derived from extensive observations, often over several years. In all instances, these observations were made by mental health professionals on the bases of criteria independent of PORT scores. The original observation protocol is (briefly) summarized.

Emphasis was on (mainly nonverbal) ways a child demonstrated comfort and effectiveness during or immediately subsequent to interactions with each caretaker. These interactions took place in spontaneous, structured, and instructional contexts. The basic dimensions were a child's movements: toward (positive); against (aggressive); away from (fearful) (Bricklin, B. & Halbert, M. 2004a, b). (See also Bargh, J.A. The automaticity of everyday life. In R.S. Wyer, Jr. (Ed.), *The automaticity of everyday life: Advances in social cognition*. 1997, Vol. 10, pp. 1-61; Duckworth, K.L., Bargh, J.A., Garcia, M. & Chaiken, S. The automatic evaluation of novel stimuli. *Psychological Science*, 2002, 13(6), pp. 513-519; Horney, K. *Our inner conflicts*. 1945; Zajonc, R.B. Feeling and thinking: Preferences need no inferences. *American Psychologist*, 1980, 35, 151-175; Zajonc, R.B. Emotions. In D.T. Gilbert, S.T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology*. 4th Ed. 1998, Vol. 1, pp. 591-632.) Research interest was on what impact parental behavior had on a child, not

primarily on what parents knew or did. At the outset of our research we tracked positive *and* negative reactions. Our research goals were adequately achieved if we counted only positive reactions. This is a practical “plus,” since fewer raters are needed. The following categories apply when the child is speaking (initiating or responding), listening, or listening and acting: smooth breathing patterns; body movements nonhurried; relaxed and/or smiling facial muscles (no grimaces, contortions, etc.); leans toward other person; maintains reasonable eye contact (evaluators were taught to distinguish fearful eye contact from relaxed eye contact---the former is motivated by the child’s fear of saying or doing something the parent would object to); moves closer to or initiates physical contact; willing to be hugged; few signs of restlessness (even if a child is ADHD-like, one parent usually has a more calming influence than the other). Categories used when the child is speaking, initiating, or responding: willing to express annoyance, doubts, or confusion (*not* trying to be the perfect little child); pauses without fear of losing caretaker’s attention; willing to ask questions; noting from whom the child most frequently and spontaneously seeks help. Categories used when the child is listening or listening and acting: accepts limits in comfortable manner; muscularly comfortable with failures (no agitated moves); focused attention; facial expression animated and interested; no evidence of leaping-to-action, i.e., the child waits for the entire “sent message” to be received; open and ready to receive information; willing to explore and take chances; willing to try novel approaches. More statistical data on the observation protocol is given in Bricklin and Halbert (2004a, b). Briefly, in two samples ($n=60$; $n=37$), the ages, ethnicity, and SES’s were directly similar to the norms given in Bricklin and Halbert, 2004. In each sample, a child was observed with both parents present, so he or she could make choices about interactions.

In a one-hour session, the distribution of positive scores ranged from 0 to 12. The mean number of positive reactions was 7.4; the Standard Deviation was 1.2. Six to nine positive reactions characterized 70 percent of the cases. A point difference greater than two represents a significant difference between two caretakers. These results are *not* comparable to protocols that count the number of positive and negative interactions *initiated by parents* (Lahey, B.B., Conger, R.D., Atkeson, B.M., & Trieber, F.A. Parenting behavior and emotional status of physically abusive mothers. *Journal of Consulting and Clinical Psychology*, 1984, 52, pp. 1062-1071; Kerig, P.K. & Lindahl, K.M. (Eds.) *Family observational coding systems: Resources for systemic research*. 2002). Interrater agreement (three raters) was high, 90 percent, partly because the range of categories used was narrow: A>B; B>A; A=B; neither A nor B. The actual number of positive interactions noted by each rater achieved an agreement rate of 82 percent. Note that while 7.4 positive reactions in a one-hour session may seem like a small number, the fact is that each involved child was often concerned with task-mastery behaviors during the session, not just in interacting with parents.

PORT validity data collected between 1961 and 1997 consisted of 1381 cases. The sources of independent validity designations involved: (1) three observers watching each child and two parents interacting from behind a one-way screen; (2) two psychologists who had access to family therapy notes and consultations with the family therapists over a two- to five-year time span; (3) scores derived from the BPS; and (4) psychologists' findings based on all clinical and life history data available to them, usually gathered over a multimonth period. The average percent of agreement between PORT suggestions for parent-of-choice or POC and those of the independent experts was 88 percent. The agreement rate between courtroom judges and the POCs selected by the PORT was 92 percent, although this is reported as information only and not as validity data, since PORT data may have influenced the judges. PORT validity data collected between 1997 and 2002 involved 127 cases. The independent validity experts were mental health professionals who had had at least eight months of contact with each involved family. They were instructed to use all sources of information including the observational format already described. Future validity was measured by comparing the PORT POCs at Time Point 1 with expert opinion carried out at Time Point 2, eight months later. Future validity was 89 percent agreement. Concurrent validity was estimated by comparing the PORT POC at Time Point 2 with the validity designations made at Time Point 2. This figure was 91 percent. BPS validation from 1964 to 1997 was available for 2279 cases. The same validation sources were used as with the PORT, only here the family therapy data extended from two to seven years (the validity designations were never based on less than two years' worth of data). The agreement rate between the BPS and the criterion experts was 88 percent. The agreement rate for courtroom judges was 93 percent. BPS validity data collected from 1997 to 2002 on 93 cases showed a future validity score of 87 percent and a concurrent validity score of 91 percent.

Test-retest reliability will be given only from the most recent data, since the time span was longer than those previously used. Between 1997 and 2002, on 127 cases, the PORT showed a stability in POC, over an eight-month span, of 97 percent. That is, 97 percent of the POCs remained stable over this time span. However, the test-retest reliability drops sharply as the so-called Task Difference Score approaches 0 or 1 (a 21 percent chance that the POC will shift over an eight-month period). If the TDS is three or more there is a 3 percent chance of shift. With the BPS, if the item difference score is 0, 1, 2, or 3 there is a 19 percent chance the POC will shift over eight months. If the score is 4 or more, there is a 3 percent chance of shift.

PORT normative data gathered from 1961 to 1997 involved 797 girls and 784 boys. The mean age was 7.76, the SD 0.17. The SES was low to high middle. There were 98 percent Caucasian in this sample and 2 percent all others. On data gathered between 1997 and 2002, there were 61 girls and 66 boys. The other numbers are all similar, except there were 8 percent non-Caucasian. For the BPS, on 2389 cases between 1964 and 1997, there were 1202 girls and

1187 boys. The mean age of the BPS sample was 8.94, the SD 2.40. All other data were similar to the PORT. Interrater reliability of PORT scoring was obtained from two samples of seminar attendees ($n = 36$; $n = 41$), in which more than half of the scorers had no prior experience with the PORT. Four different percent-of-agreement scores were obtained: (1) the points scored on Task I (the most complex task); (2) the POC on Task I; (3) the overall Task Difference Score for all seven tasks; (4) the overall POC based on seven tasks. The percent-of-agreement rates, respectively, were: 74; 90; 82; 92. No interrater data for the BPS were gathered, since scoring it is mechanical and requires only the ability to read Arabic numbers and to recognize when one is larger than another. It is also assumed that an evaluator can add and subtract numbers between zero and 32.

The next research is the first in which PORT data were used with a group reference (Bricklin, 2004, Custody Newsletter). A consistent relation between PORT signs and the adequacy of parenting was noted from PORTs collected over a 40-year span. Twenty-three hypotheses about them were developed. Following this, from the huge pools of data available, four experimental groups matched in age, income, absence of ADHD signs and age-correct grade placement were formed. (Subsequent research showed conclusively that ADHD does not create bias in the PORT). Group I ($n = 16$) were children examined for minor reasons, mostly underachievement; none involved the quality of parenting. Group II ($n = 34$) were children whose parents were involved in mild custody disputes, centered on who could provide a better school, neighborhood, or extended family, or an argument between the parents based on the claim by each that the involved child had more friends at one place rather than the other. There were no concerns about parenting by either side or the evaluators. Group III ($n = 40$) consisted of children whose parents were involved in continual conflict, characterized by ongoing hostility, often within the courtroom, for two or more years. The involved children, "caught in the middle," were used as "message carriers" to deliver scathing messages from one parent to the other or to involved mental health professionals. Parental adequacy sinks to a serious low point when parents become so engaged, and the damage to children, whether from divorced or intact families, is alarmingly high (Hoppe, C.F. Test characteristics of custody-visitation litigants: A data-based description of relationship disorders. In S. Podrygula (Chair), *Empirical approaches to child custody determination*. Symposium conducted at the meeting of the American Psychological Association, August, 1993, Toronto, Ontario, Canada; Hoppe, C.F. Perpetually battling parents. In B. Mark & J. Incoravia (Eds.), *The handbook of infant/child and adolescent psychotherapy, Vol. II*, 1997, pp. 485-501; Hoppe, C.F. & Kenney, L. *Characteristics of custody litigants: Data from the Southern California group*. Paper presented at the meeting of the Psychological Association, August, 1994, Los Angeles, CA; Hoppe, C.F. & Kenney, L. *MMPI-2 and Rorschach profiles of custody litigants: An intercorrelational study*. Paper presented at the meeting of the American Psychological Association, August, 1995, New York; Hoppe, C.F.

& Kenney, L. *Therapeutic intervention in high conflict divorce: Countertransference and the horrible decision*. Paper presented at the meeting of the American Psychological Association, 1997, Chicago, IL; Bricklin, B. & Elliot, G. *ACCESS: A comprehensive custody evaluation standard system*. 1995, pp. 38-40; Bricklin, B. & Elliot, G. Qualifications of and techniques to be used by judges, attorneys and mental health professionals who deal with children in high conflict divorce cases. *University of Arkansas Little Rock Law Review*, 2000, 122, pp. 501-505; Doolittle, D.B. & Deutsch, R. Children and high conflict divorce: Theory, research, and intervention. In R.M.Galatzer-Levy & L. Kraus (Eds.) *The scientific basis of child custody decisions*, 1999, pp. 425-440). Group IV ($n = 40$) consists of children whose caretakers represented on the PORTs had either been threatened by the court with a possible termination of parental rights or actually had had their parental rights terminated in the past and/or were accused of substantiated abuse or neglect. In order to match the socioeconomic status of the parents in this group, and to match the intelligence of these children to those in other groups, the majority of these cases were derived from various private practices.

Seventeen of the 23 PORT signs yielded a probability of .05 or less so that the differential occurrences of the PORT signs among the groups could be random. Statistical effect sizes were substantial. Several caveats are listed to show that the manifestation of one or even several of these signs should be interpreted currently only as red-flags to launch expanded evaluations, since it is not completely clear that the signs are additive, nor is the interdependence among them known.

Normative data are developed because in almost all tests the respondents' answers derive (statistical) meaning by how any given individual's scores compare to those of previously examined groups. The so-called "norms" allow a test user to know how any tested person matches the people in these databases. As of now, there are published databases of 1,508 cases for the PORT and 2,372 for the BPS.

2. Has the method been published and peer-reviewed? The answer is yes to both questions. Please note that many of the reviews were written before the databases were anywhere near 4,000 cases. Note also that of those reviews that included negative commentary, *all* were based on theoretical issues. No empirical research data has ever been brought forward to challenge the accuracy statistics of either the PORT or BPS.

3. Does the method yield a known error rate? The answer is yes. Accuracy or validity figures between test hypotheses and results obtained range from about 88 to 92 percent.

4. Has the method found "general acceptance" in the field to which it is relevant? The answer is yes. In a 1997 study (Ackerman, M.J. & Ackerman,

M.C. Custody evaluation practices: A survey of experienced professionals (revisited). *Professional Psychology: Research and Practice*, 1997, 28(2), pp. 137-145) the two tests were the most widely used custody-related tests for children in the country.

5. Have the author(s) of the tests provided ongoing data about ways to maximize the tests' reliability? The answer is yes. Countless updates to the manuals have been published and made available to test users. Great emphasis has been placed on helping a mental health professional recognize a child who is deliberately "campaigning" for a particular parent not based on his or her actual interactions with that parent, but rather by being motivated by fear, intimidation, bribery, lies, the desire to avenge a parent seen as hurt, or the wish to "save" a parent seen as hurt or impaired. A group-referenced study helps evaluators to red-flag serious parental pathology.

SUMMARY: THE DAUBERT DICTA AND THE PORT AND BPS

1. Are the PORT and BPS "testable," and have they been tested? The answer is yes to both questions. The combined database is 3,880.

2. Have the tests been published and peer-reviewed? The answer is yes.

3. Is there a known error rate? Yes; the predictive accuracy of the tests ranges between 88 and 92 percent agreement between test-based assertions and a wide range of validating criteria.

4. Have the tests found "general acceptance" in the mental health field? The answer is yes. A study found them to be the most widely used custody tests for children in the country.

5. Have the authors provided test users with information helpful in maximizing the validity and reliability of the tests? The answer is yes, in dozens of Manual Updates and in the current Manual revision.

SOME PERSONAL TIPS TO THE MENTAL HEALTH PROFESSIONAL GOING TO COURT

1. This paper approaches admissibility issues from a national perspective. Make sure you find out what the legal criteria and guidelines pertaining to admissibility of evidence as well as of expert testimony are in your state.

2. Make sure that any attorney who represents your interests is up-to-speed on admissibility issues. If he or she is not, a courtroom judge will not come to your rescue.

REFERENCES

- Ackerman, M.J. & Ackerman, M.C. (1997). Custody evaluation practices: A survey of experienced professional (revisited). *Professional Psychology: Research and Practice*, 28(2), 137-145.
- Ackoff, R.L. (1999). *Ackoff's Best*. New York: Wiley
- Bargh, J.A. (1997). The automaticity of everyday life. In R.S. Wyer, Jr. (Ed.) *The automaticity of everyday life; Advances in social cognition* (Vol. 10, pp. 1-61). Mahwah, NJ: Erlbaum.
- Bricklin, B. (1984). *The Bricklin Perceptual Scales*. Furlong, PA: Village Publishing.
- Bricklin, B. (1989). *Perception-of-Relationships Test*. Furlong, PA: Village Publishing.
- Bricklin, B. (1992). Data based tests in custody evaluations. *American Journal of Family Therapy*, 20, 254-265.
- Bricklin, B. (1995). *The custody evaluation handbook*. New York: Brunner/Mazel.
- Bricklin, B. (1997). *Test manuals supplement update*. Furlong, PA: Village Publishing.
- Bricklin, B. (1998). Sequence of steps and critical assessment targets of a comprehensive custody evaluation (pp. 499-502). In G.P. Koocher, J.C. Norcross, & S.S. Hill III (Eds.), *Psychologists' desk reference*. New York: Oxford University Press.
- Bricklin, B. (1999). The contribution of psychological tests to custody-relevant evaluations (pp. 120-156). In R.M. Galatzer-Levy & L. Kraus (Eds.), *The scientific basis of child custody decisions*. New York: Wiley.
- Bricklin, B. (2002). *The frequency of appearance of twenty-three PORT signs among four groups representing different degrees of parental competence*. Furlong, PA: Village Publishing.
- Bricklin, B. (2004). Parental value and adequacy as reflected in the Perception-of-Relationships Test: Empirical data on seventeen statistically significant signs. *The Custody Newsletter, Issue 21*, 77-126.

- Bricklin, B. & Bricklin, P.M. (1999). Custody data as decision-theory information: Evaluating a psychological contribution by its value to a decision-maker. *Clinical Psychology: Science and Practice*, 6(3), 339-343.
- Bricklin, B. & Elliot, G. (1995). *ACCESS: A comprehensive custody evaluation standard system*. Furlong, PA: Village Publishing.
- Bricklin, B. & Elliot, G. (2000). Qualifications of and techniques to be used by judges, attorneys and mental health professionals who deal with children in high conflict divorce cases. *University of Arkansas Little Rock Law Review*, 122, 501-528.
- Bricklin, B. & Elliot, B. (2001). *The Bricklin family of custody instruments: Updated validity, reliability and normative data*. Furlong, PA: Village Publishing.
- Bricklin, B. & Elliot, G. (2002a). What can empirical data on 4,500 child custody cases tell us? Paper presented at the American College of Forensic Psychology, San Francisco, CA, April 18, 2002.
- Bricklin, B. & Elliot, G. (2002b). *The Perception-of-Relationships Test (PORT) and Bricklin Perceptual Scales (BPS)*. Furlong, PA: Village Publishing.
- Bricklin, B. & Elliot, G. (2005). Empirically assisted assessment of family systems. In L. Gunsberg & P. Hymowitz (Eds.), *A Handbook of Divorce and Custody: Forensic, Developmental, and Clinical Perspectives* (pp. 201-219). Hillsdale, NJ: The Analytic Press.
- Bricklin, B. & Halbert, M.H. (2004a). Can child custody data be generated scientifically? Part I. *The American Journal of Family Therapy*, 32, 119-138.
- Bricklin, B. & Halbert, M.H. (2004b). Perception-of-Relationships Test and Bricklin Perceptual Scales: Validity and reliability issues. Part II. *The American Journal of Family Therapy*, 32, 189-203.
- Doolittle, D.B. & Deutsch, R. (1999). Children and high-conflict divorce: Theory, research, and intervention. In R.M. Galatzer-Levy & L. Kraus (Eds.) *The Scientific Basis of Child Custody Decisions* (pp. 425-440). New York: John Wiley & Sons, Inc.
- Duckworth, K.L., Bargh, J.A., Garcia, M. & Chaiken, S. (2002). The automatic evaluation of novel stimuli. *Psychological Science*, 13(6), 513-519.
- Einstein, A. (1936). Physics and reality. *Journal of the Franklin Institute*, 221(3), 1-34.

- Einstein, A. (1984). *Out of my later years*. New York: Kensington.
- Groscup, J.L., Penrod, S.D., Studebaker, C.A., Huss, M.T., & O'Neil, K.M. (2002). The effects of Daubert on the admissibility of expert testimony in State and Federal criminal cases. *Psychology, Public Policy and Law*, 8(4), 339-372.
- Hoppe, C.F. (1993). Test characteristics of custody-visitation litigants: A data-based description of relationship disorders. In S. Podrygula (Chair) *Empirical approaches to child custody determination*. Symposium conducted at the meeting of the American Psychological Association, Toronto, Ontario, Canada, August.
- Hoppe, C.F. (1997). Perpetually battling parents. In B. Mark & J. Ingoravia (Eds.), *The handbook of infant/child and adolescent psychotherapy*, Vol. 2 (pp. 485-501). Northvale, NJ: Aronson.
- Hoppe, C.F. & Kenney, L. (1994). Characteristics of custody litigants: Data from the southern California group. Presented at the meeting of the American Psychological Association, Los Angeles, CA, August.
- Hoppe, C.F. & Kenney, L. (1995). MMPI-2 and Rorschach profiles of custody litigants: An intercorrelational study. Presented at the meeting of the American Psychological Association, New York, NY, August.
- Hoppe, C.F. & Kenney, L. (1997). Therapeutic intervention in high conflict divorce: Countertransference and the horrible decision. Presented at the meeting of the American Psychological Association, Chicago, IL, August.
- Horney, K. (1945). *Our inner conflicts*. New York, NY: Norton.
- Kerig, P.K. & Lindahl, K.M. (Eds.) (2001). *Family observational coding systems*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Larson, E. (2006). *Thunderstruck* (pp. 17 and 312). New York: Crown Publishing Group.
- Lahey, B.B., Conger, R.D., Atkeson, B.M., & Trieber, F.A. (1984). Parenting behavior and emotional status of physically abusive mothers. *Journal of Consulting and Clinical Psychology*, 52, 1062-1071.
- Lipton, J.P. (1999). The use and acceptance of social science evidence in business litigation after Daubert. *Psychology, Public Policy and Law*, 5(1), 59-77.

Mark, M.M. (1999). Social science evidence in the courtroom: Daubert and beyond. *Psychology, Public Policy and Law*, 5(1), 175-193.

Popper, K. (1959). *The logic of scientific discovery*. Original edition 1934. Springer: Vienna, Amplified English Edition.

Sales, B.D. & Shuman, D.W. (1999). The impact of Daubert and its progeny on the admissibility of behavioral and social science evidence. *Psychology, Public Policy and Law*, 5(1), 3-15.

Segre, G. (2007). *Faust in Copenhagen: A struggle for the soul of physics*. New York: Viking Penguin.

Weiner, I.B. (2001). Assessment advocacy. *Social Personality Assessment Exchange*, 12(1), 7.

Zajonc, R.B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35, 151-175.

Zajonc, R.B. (1998). Emotions. In D.T. Gilbert, S.T. Fiske & G. Lindzey (Eds.), *Handbook of social psychology* (4th Edition, Vol. 1, pp. 591-632). New York: Oxford University Press.