Use of the Personality Assessment Inventory in Child-custody Evaluation

Daniel J. Hynan, Independent Practice, Naperville, IL
Email: drdanhynan@gmail.com

Abstract

Child-custody evaluations have been criticized for having an inadequate scientific basis (e.g., Emery, Otto, & O'Donohue, 2005). Objective psychological testing with valid and reliable instruments, which have child-custody reference data, can strengthen the scientific foundations of such evaluations. Personality Assessment Inventory (PAI) (Morey, 2007) scores based on 250 child-custody litigants are described. Moderate defensive underreporting was exhibited. Seven scales had statistically significant gender differences. The mean for fathers on the mania scale was higher than for mothers, which is inconsistent with epidemiological findings on gender prevalence. Otherwise, results suggest that the PAI has adequate gender fairness. These data are accompanied by interpretive considerations associated with the custody evaluation context and the importance of integrating test results with other evaluative information.

Keywords: child custody, defensiveness, evaluation, gender, personality

A great deal of research, covering tens of thousands of individuals, indicates that children of divorce function less well than those from never-divorced families in numerous areas of life (Amato & Keith, 1991a). However, it is clear that not all children who go through the divorce of their parents experience a significant negative effect. Research has investigated a number of factors that might contribute to negative outcomes for children of divorce. One of the factors consistently identified as a significant risk for a negative outcome is parental psychological dysfunction that impacts parenting behavior (Amato, 1993, 2001; Amato & Keith, 1991a, 1991b; Hetherington, Bridges, & Insabella, 1998; Lansford, 2009).

In addition, many state statutes make reference to psychological functioning as a relevant factor in child-custody determinations. The Uniform Marriage and Divorce Act (UMDA) (National Conference of Commissioners on Uniform State Laws, 1979) is relied upon, at least in part, by many statutes that pertain to child custody (Elrod & Dale, 2008-2009). The UMDA indicates that the mental health of all relevant parties is a factor to be considered in custody determination. According to the Specialty Guidelines for Forensic Psychology (American Psychological Association (APA), 2013), forensic examiners need to include a focus on legally relevant factors in their assessments.

In child-custody evaluations, psychological testing is routinely used. A frequent objective of testing adults is to clarify the possible presence of psychological maladjustment. It is important to use well-validated psychological tests to increase reliance on scientific

Hynan, D. J. (2013). Use of the Personality Assessment Inventory in child custody evaluation. *Open Access Journal of Forensic Psychology, 5,* 120-133.

data, and thereby augment more subjective clinical impressions, the latter of which may be more likely to reflect judgment errors (Garb, 1998).

It is crucial for test usage to be relevant for the testing objective. Numerous professional principles and guidelines address the importance of appropriate and fair psychological test use. For example, the APA (2002) Ethical Principles of Psychologists and Code of Conduct encourage test use that is valid and reliable for the purpose of the assessment and the test taker. The Standards for Educational and Psychological Testing (American Educational Research Association, APA, & National Council on Measurement in Education, 1999) address numerous areas, including interpretation informed by appropriate reference groups. The Specialty Guidelines for Forensic Psychology (APA, 2013) state that test use needs to consider applicable research. Similarly, the Guidelines for Child Custody Evaluations in Family Law Proceedings (APA, 2010) indicate that interpretation of assessment data needs to consider the context of the child-custody evaluation. In addition, all these documents emphasize the importance of fairness and avoiding unwarranted bias.

The PAI is an omnibus objective measure of psychological disorders, personality functioning, and related features. It is the third most frequently used measure of that type in custody evaluation (Ackerman & Pritzl, 2011), but there are no published data specific for such use. Regarding tests commonly used in custody evaluation, there are published custody evaluation data that include gender comparisons only for the Minnesota Multiphasic Personality Inventory-2 (MMPI-2) (Bagby, Nicholson, Buis, Radovanovic, & Fidler, 1999; Bathurst, Gottfried, & Gottfried, 1997; Strong, Greene, Hoppe, Johnston, & Olesen, 1999) and the Millon Clinical Multiaxial Inventory-III (MCMI-III) (McCann et al., 2001), though the latter measure has been questioned in terms of its validity (Rogers, Salekin, & Sewell, 1999) and gender fairness (Hynan, 2004). Other custody-evaluation data from the MMPI-2 (Ezzo, Pinsoneault, & Evans, 2007; Posthuma & Harper, 1998; Siegel, 1996; Siegel & Langford, 1998) and NEO Personality Inventory (Langer, 2011) do not include useful gender comparisons for all main scales.

The PAI has eleven scales of clinical functioning and two of interpersonal style relevant to custody evaluation. In addition, there are four validity scales and five that pertain to treatment considerations. All main scales are listed in Table 1. The Positive Impression Management (PIM) validity scale, the main measure of defensiveness, is especially important because evaluators need to consider strongly the extent to which parents are trying to present themselves in the best possible light, and such clients have been found to manifest moderately elevated levels on the MMPI-2 (Bagby et al., 1999; Bathurst et al., 1997; Posthuma & Harper, 1998; Siegel, 1996; Strong et al., 1999).

Table 1

Descriptive Statistics of the PAI Scales for the Sample of Child-Custody Litigants

	Total (n = 250)		Mothers (n = 125)		Fathers (n = 125)	
Scale	*M*	*SD*	*M*	*SD*	*M*	*SD*
Validity						
ICN	44.25	6.53	44.27	7.15	44.22	5.88
INF	50.22	7.74	50.31	7.75	50.14	7.76
NIM	45.57	3.33	45.59	3.83	45.54	2.76
PIM	60.24	7.97	60.60	7.96	59.88	8.00
Clinical						
SOM	44.28	5.12	44.33	4.16	44.24	5.96
ANX	42.74	5.64	43.53	6.04	41.94	5.13
ARD	40.16	6.81	40.19	7.00	40.12	6.65
DEP	42.45	5.15	42.63	5.11	42.26	5.20
MAN	43.44	8.12	41.42	7.48	45.46	8.26
PAR	44.38	8.47	44.15	8.54	44.60	8.43
SCZ	40.00	5.53	39.90	5.78	40.10	5.28
BOR	42.64	5.71	42.62	5.85	42.66	5.59
ANT	44.10	5.95	42.60	5.02	45.61	6.43
ALC	45.21	4.92	44.44	4.55	45.98	5.17
DRG	46.04	5.51	46.34	5.29	45.74	5.73
Treatment						
AGG	41.60	5.81	41.17	5.37	42.03	6.20
SUI	44.96	3.94	45.07	3.39	44.86	4.43
STR	52.96	10.29	52.94	9.91	52.98	10.69
NON	40.90	5.05	40.57	4.98	41.24	5.11
RXR	57.66	7.22	56.52	6.98	58.81	7.31
Interpersonal						
DOM	53.26	7.37	51.89	7.60	54.63	6.89
WRM	58.04	7.12	60.02	6.45	56.06	7.22

Note. ICN = Inconsistency; INF = Infrequency; NIM = Negative Impression; PIM = Positive Impression; SOM = Somatic Complaints; ANX = Anxiety; ARD = Anxiety-Related Disorders; DEP = Depression; MAN = Mania; PAR = Paranoia; SCZ = Schizophrenia; BOR = Borderline Features; ANT = Antisocial Features; ALC = Alcohol Problems; DRG = Drug Problems; AGG = Aggression; SUI = Suicidal Ideation; STR = Stress; NON = Nonsupport; RXR = Treatment Rejection; DOM = Dominance; WRM = Warmth

There is good evidence for the reliability and validity of the PAI. For example, except for two validity scales (ICN and INF), the internal consistency and test-retest reliability are adequate to good, and there is considerable evidence of concurrent validity (Morey, 2007). Additional evidence of validity can be seen in clinical and offender samples (e.g., Edens, Hart, Johnson, Johnson, & Olver, 2000; Jacobo, Blais, Baity, & Harley, 2007).

One main objective of the current research was to provide relevant data for child-custody evaluators. There were specific predictions, also. First, it was expected that the results would show, compared to the standardization sample (Morey, 2007), a moderate elevation in PIM and general decreases in clinical-scale elevations, based on informal observation and MMPI-2 findings (e.g., Bagby et al., 1999; Bathurst et al., 1997). Second, informal observation led to a prediction that interpersonal scales, especially WRM, would show modest elevations.

Third, it was predicted that the PAI would manifest little to no difference between genders on scale elevations. Gender differences are extremely important because it is crucial to have fairness and avoid unwarranted bias, and child custody evaluations almost always involve comparisons between women and men. Prior research found that women more frequently obtained higher MCMI-III clinical Base-Rate scores than men on scales that measured histrionic, narcissistic, and obsessive-compulsive personality, despite the fact that there were essentially no gender differences in raw scores, and the Base-Rate score gender differences were entirely different from the results of epidemiological research on the prevalence of such disorders (Hynan, 2004). Therefore, it is crucial to look for possible gender differences in test scores as a first step in considering whether a measure has adequate gender fairness when used in custody evaluation.

Method

Participants

Archival data from the author's practice in suburban Chicago served as the database. All 250 participants were legal parents. That number was far above the requirement for a conventional power level of .80 (meaning an 80% probability of rejecting a false null hypothesis) given a statistical significance level of .05 and a goal of identifying at least a medium effect size (Cohen, 1992). Parental demographic data were based on information completed on the PAI answer sheet. Mothers averaged 37 years of age (*SD* = 6.09) and fathers 40 (*SD* = 7.51). Mothers and fathers both had an average of 15 years of education (*SD* = 2.57). There was a mean of 1.72 children per family (*SD* = .75). The vast majority of participants were Caucasian.

Materials and Procedures

All tests were completed as a part of evaluations ordered by court. Twenty-seven tests were completed in accordance with a local court rule for a condensed evaluation, and the rest were completed according to standard orders without any such limitation. All tests from 1998 through 2009 were used in this analysis. Almost all tests had no unanswered items, and none had more than 17 unanswered, the maximum allowed for interpretation (Morey, 2007). The author administered all tests in booklet form in the office setting. All were computer scored by Psychological Assessment Resources.

Results

Descriptive statistics (T-scores) are displayed in Table 1. Consistent with predictions, PIM was moderately elevated and clinical-scale mean elevations were lower than the PAI standardization sample. Treatment-consideration-scale elevations were also lower, except for STR and RXR. Although it was not included in the research hypotheses, it is noteworthy that RXR was moderately correlated with PIM (r = .54, p < .001); WRM was mildly elevated and correlated with PIM (r = .29, p. < .001)

To test for gender differences, a multivariate analysis of variance was conducted with gender as the between-subjects factor and the 22 primary scales as the multiple dependent measures. Results showed a significant interaction effect of gender by scale, multivariate F (21, 228) = 4.91, p <.001.

Gender differences were examined by means of an ANOVA carried out for each of the scales. See Table 1 for means and standard deviations. Women had higher scores on ANX (F (1, 248) = 5.00, p = .026 and WRM (F (1, 248) = 21.00, p < .001. Men had higher scores on MAN (F (1, 248) = 16.51, p < .001; ANT (F (1, 248) = 17.00, p < .001; ALC (F (1, 248) = 6.29, p = .013; RXR (F (1, 248) = 6.41, p = .012; and DOM (F (1, 248) = 8.94, p = .003. There were no statistically significant differences on other scales. It should be noted that the PAI does not have gender differences in how raw scores are translated into T-scores.

For most of the scales that showed statistically significant gender differences, the mean differences were small and within the standard error of measurement of 4 T-score units; mean differences within the standard error of measurement are unlikely to be clinically significant (Morey, 2007). Among the clinical scales, the mean gender difference for MAN was 4.02, with fathers showing the higher score. Women had a mean score on WRM, an interpersonal scale, which was 3.96 higher than men and a difference almost identical to the standard error of measurement.

Another method of considering clinical significance is the effect size of statistically significant differences. On clinical scales, the higher scores for fathers on MAN and ANT showed medium effect sizes (Cohen's (1992) ds = 0.51 and 0.52, respectively), whereas the higher score on ALC had a small effect size (d = 0.32). The higher score

for women on ANX had a small effect size (*d* = 0.28). On interpersonal scales, the higher score for mothers on WRM was at a medium effect size (*d* = 0.58), and the higher score for men on DOM was at a small effect size (*d* = 0.38). Although the treatment consideration scales are seldom very relevant within the context of custody evaluation, it should be noted that the RXR higher score for males manifested a small effect size (*d* = 0.32).

Discussion

These test results provide the first PAI comparison scores for use by custody evaluators, and a potential starting point for future research using this measure within the child-custody evaluation context. Consistent with predictions, there was a moderate elevation in PIM, equivalent to approximately one standard deviation. Also consistent with predictions, all clinical scales had mean elevations lower than the standardization sample. The mean clinical scale elevations were approximately one-half to one standard deviation lower than the PAI standardization sample. The prediction that there would be a modest elevation in the mean interpersonal scale elevation was largely accurate, although the mean elevation for WRM was almost one standard deviation above the mean for the standardization sample.

There was partial support for the third prediction regarding gender differences. Most scales showed no statistically significant gender differences. Of those that manifested significant gender differences, most showed small effect sizes and therefore are unlikely to make a substantive difference in custody evaluation practice. However, contrary to predictions, the clinical scales MAN and ANT and the interpersonal scale WRM manifested gender differences that had medium effect sizes.

It is useful to compare the current findings regarding the PIM defensiveness scale with empirical findings from the MMPI-2 defensiveness scales *L* and *K* when used with custody evaluation parents. Because both the PAI and the MMPI-2 use T-scores, the elevations can be directly compared. The three MMPI-2 custody evaluation studies that had the largest samples (Bagby et al., 1999; Bathurst et al., 1997; Strong et al., 1999) reported *L* and *K* scores between 57 and 61. These scores are very similar to the current PIM mean of 60, especially considering that the standard deviations for all these defensiveness scale scores are approximately in the range of 8 to 11. Overall, it is not possible to know to what extent the defensiveness found here and in past research represents relatively healthy individuals completing tests in a defensive manner, and/or to what extent the defensiveness may serve to cover up significant psychopathology in a number of the participants.

It is likely that the general moderate defensiveness found in the current PAI sample contributed to the generally lowered mean scores on all clinical scales. It is important to note that the PAI does not adjust clinical scale scores in a manner similar to the MMPI-2 scale *K* correction. The larger MMPI-2 studies that used the *K* correction (Bagby et al., 1999; Bathurst et al., 1997) generally found mean scores on the main clinical scales

that were closer to the standardization sample mean, and a number of those scores were at or above the mean. Although a reasonable hypothesis would be that the *K* correction was the reason that the MMPI-2 studies did not show the same extent of decrease in clinical scale scores as the current PAI results, an examination of the data suggests that is not the case. That is, some of the relatively elevated MMPI-2 scales in the Bagby and Bathurst samples, such as Pa, are not *K* corrected, and some that showed below-average scores, such as Sc, are *K* corrected. It is difficult to know the reasons for the somewhat different patterns of results between the current research and those past MMPI-2 studies, because there are many differences between the PAI and MMPI-2. A possible factor is that these differences are due in part to the different samples, and it is noteworthy that there are a number of specific, if modest, differences in the patterns of results between the Bagby and Bathurst findings.

The fact that WRM showed an elevation of just under one standard deviation, and had a modest positive correlation with PIM, may have particular relevance for evaluators. The PAI manual (Morey, 2007) states that the scale measures interpersonal features such as empathy, patience, affection, and sociability. All are likely to be seen as positive features in a parent. Therefore, evaluators need to exercise caution in making interpretations about elevated WRM scores, especially given that at least some aspect of a significant elevation may be associated with the parent presenting him/herself in a positive light. The mean score for DOM is only a very small amount higher than the standardization sample and therefore is likely not of practical importance. Elevated scores on RXR are likely associated with defensive responding in many cases.

Gender differences on clinical scales found here reflect prevalence differences found for relevant clinical conditions, except for MAN. More specifically, the fact that there was a significantly higher score for women than men on ANX is representative of the fact that women generally manifest a higher prevalence of anxiety disorders than do men. On ANT and ALC, there was a significantly higher score for men than women. These mean score differences reflect the fact men generally show higher prevalence rates of antisocial personality and alcohol misuse than do women (Eaton et al., 2012; World Health Organization, 2007).

On MAN, there was a significantly higher mean score for fathers than mothers, with a medium effect size. However, epidemiological research has not found a difference between genders in the prevalence of bipolar disorders (Ferrari, Baxter, & Whiteford, 2011; World Health Organization, 2007). Therefore, there is evidence that this scale may be unfairly weighted against fathers within the context of a child-custody evaluation.

However, it is important to note the relatively low mean scores on this scale for both women and men in the current sample. When both the mean and standard deviation for fathers on MAN are considered, a male who scores even two standard deviations above average would have a score of approximately 62, which is a modest elevation when compared to the standardization sample. The PAI manual states, within that range, a

respondent may show positive features such as being active and self-confident, though might also "be rather impatient and hostile, with a quick temper" (Morey, 2007, p. 38). The manual describes two higher MAN score ranges, starting at T-scores of 65 and 75, which represent more severe levels of dysfunction.

Also, it is important to note the very considerable differences between the empirically unsupported disadvantage for males on MAN and the same type of disadvantage for women on three different personality disorder scales of the MCMI-III. On that latter measure, a considerable proportion of women were found to have scores representative of personality disorders or maladaptive personality traits (Hynan, 2004), whereas on the PAI, only an extremely small proportion of males appear likely to have significant elevations on MAN.

When such an elevation is found, it would be very important to look for other evidence in the evaluation to either support or contradict the likely presence of manic-type behaviors. Of course, such an examination of the overall record for confirmatory or disconfirmatory evidence needs to be a routine evaluation practice regarding various types of data collection.

The PAI standardization sample (Morey, 2007) also shows some gender differences. For MAN, males have a higher mean score than do females, but the mean difference and effect size are smaller than in the current custody evaluation sample. For ANT and ALC, males also have higher mean scores than do females, but the differences and effect sizes are larger than in the current sample. Although women have a higher mean score on ANX than do men in the standardization sample, the difference is negligible. Regarding the interpersonal scales, in the standardization sample women had higher mean scores on WRM, though at a smaller effect size than the current sample. On DOM, the standardization sample results are similar to the current sample in that men have a higher mean score, and the effect size is almost identical. There is essentially no difference between genders in the standardization sample on RXR. These effect size estimates are based on Cohen's *d* (Cohen, 1992) and were carried out by the author using data provided in the PAI manual (Morey, 2007, pp. 92-93). For evaluators, the main practical relevance of these comparisons with the standardization sample is the larger gender difference on MAN in the current sample.

In addition, it is important to note that, in the current sample, there are a number of scales that do not show gender differences, but research indicates that there are different gender prevalence rates. As can be seen in Table 1, there are essentially no gender differences for mean scores on SCZ and DRG, but research indicates there are higher frequencies for men of schizophrenic and drug use disorders. Although there are no gender differences for mean scores on ARD, DEP, and PAR, research reports higher frequencies for women of depressive disorders, paranoid personality disorder, and the types of anxiety disorders represented by ARD (Eaton et al., 2012; Grant et al., 2004; Longenecker et al., 2010; World Health Organization, 2007). However, the discrepancies between the lack of gender differences in mean scores in this sample and the gen-

der differences found in prevalence research are likely to be relevant for custody evaluation practice in only a very small proportion of cases, because very few parents will have significantly elevated scores on these scales.

The primary limitation of the current study is that the data were all drawn from one practice, obviously in a circumscribed geographical area. It is a limitation that there were very few non-Caucasian participants. However, it is noteworthy that such features have been present in relevant past research, also. The largest published custody evaluation psychological test sample (Bathurst et al., 1997) was also drawn from one practice. Also, there are indications that a number of demographic characteristics of the current sample are similar to prior custody evaluation MMPI-2 research. For example, the age range and level of education of parents, as well as mean number of children in this PAI sample are similar to data of prior studies, to the extent that such demographic information was available (Bagby et al., 1999; Bathurst et al., 1997; Ezzo et al., 2007; Strong et al, 1999). Such demographic similarities suggest that the current sample is not a highly unusual one, and gives reason for optimism that the results have generalizability to other geographical areas and practices. In addition, leaders in psychology have long recognized the value of collecting local norms in practice settings (Stricker & Trierweiler, 1995).

Despite the limitations of the current study, custody evaluators are likely to benefit from having a substantial reference sample for this test. In considering the balance between the usefulness and limitations of such research, it is important to note that the custody evaluation field attracts very limited attention from academic researchers, who are the individuals most likely to have the ability to collect data from multiple, disparate locations that potentially would serve as the source for more comprehensive norms. I believe that, in order to advance the custody evaluation field, scientifically minded practitioners need to take the initiative to inject a stronger empirical foundation into routine practices. It is hoped that the current study will encourage others to work toward strengthening the scientific foundations of the custody evaluation field.

Of course, the empirical data presented here need to be interpreted in a thoughtful and balanced way in order to be useful. The PAI manual (Morey, 2007) presents normative data both for a census-matched standardization sample and a clinical sample. Because there is no evidence that custody-evaluation parents generally manifest clinical problems, interpretation of test scores should primarily rely on the census-matched standardization sample. As noted above, it is important to consider the recommended interpretations based on T-score ranges described in the manual.

A major step forward in the interpretation of psychological tests of custody evaluation parents would occur if there were other sources of accurate data available about their behavioral functioning to allow external validation specific to the child custody context (Bagby et al., 1999). It is a particularly daunting challenge, however, because the defensiveness often present in test responses would also lead a significant proportion of

participants, and perhaps even collateral sources of information, to be guarded about disclosing extratest data that potentially would serve as validation criterion measures.

As noted above, in custody evaluation practice it is necessary to collect multiple sources of data. When any psychological test results suggest that an individual might have a significant dysfunction, especially one that would have a likely impact on parenting, it is crucial to look for other sources of evaluative information that might converge with or diverge from that test indication. It is also crucial to stay cognizant that psychological problems do not prevent an individual from being a good parent, and perhaps the preferred one in an evaluation in which the totality of relevant factors are fully considered.

The current research is important because parental psychological dysfunction is a factor that can have a significant negative impact on children of divorce, and psychological testing with well-validated measures such as the PAI can inject crucial scientific objectivity into the identification of such potential behavioral health problems. This study is a first step with the PAI to strengthen the empirical basis of custody evaluations, which have been criticized for excessive reliance on unscientific methods and on evaluator personal values and biases (e.g., Emery, Otto, & O'Donohue, 2005; Erickson, Lilienfeld & Vitacco, 2007). It is hoped that the current study will help to stimulate further research on this and other psychological measures for use in custody evaluation. As more well-validated tests with custody evaluation comparison data become available, it may help to reduce reliance on measures with questionable reliability and validity.

In general, strengthening the empirical foundations of custody evaluation, combined with careful professional consideration of how to apply such general scientific knowledge to the particular facts and circumstances of each case, can help to advance the field to a higher ground. Perhaps even more importantly, such enhancements in the field would be of considerable benefit to the children and families of divorce.

References

Ackerman, M. J., & Pritzl, T. B. (2011). Child custody evaluation practices: A 20-year follow-up. *Family Court Review, 49*, 618–628. doi: 10.1111/j.1744-1617.2011.01397.x

Amato, P. R. (1993). Children's adjustment to divorce: Theories, hypotheses and empirical support. *Journal of Marriage and the Family, 55*, 23–28. doi: 10.2307/352954

Amato, P. R. (2001). Children of divorce in the 1990s: An update of the Amato and Keith (1991) meta-analysis. *Journal of Family Psychology, 15*, 355-370. doi: 10.1037/0893

Amato, P. R., & Keith, B. (1991a). Parental divorce and adult well-being: A meta-analysis. *Journal of Marriage and the Family, 53*, 43–58. doi: 10.2307/353132

Amato, P. R., & Keith, B. (1991b). Parental divorce and the well-being of children: Ameta-analysis. *Psychological Bulletin, 110*, 26 –46. doi: 10.1037/0033-2909.110.1.26

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing.* Washington DC: American Educational Research Association

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57*, 1060 –1073. doi: 10.1037/0003-066X.57.12.1060

American Psychological Association. (2010). Guidelines for child custody evaluations in family law proceedings. *American Psychologist, 65,* 863–867. doi:10.1037/a0021250

American Psychological Association. (2013). Specialty guidelines for forensic psychology. *American Psychologist*, *68*, 7-19. doi: 10.1037/a0029889

Bagby, R. M.,Nicholson, R. A., Buis, T., Radovanovic, H., & Fidler, B. J. (1999). Defensive responding on the MMPI-2 in family custody and access evaluations. *Psychological Assessment, 11*, 24–28. doi: 10.1037/1040-3590.11.1.24

Bathurst, K., Gottfried, A. W., & Gottfried, A. D. (1997). Normative data for the MMPI-2 in child custody litigation. *Psychological Assessment, 9*, 205–211. doi: 10.1037/1040-3590.9.3.205

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159. doi: 10.1037/0033-2909.112.1.155

Eaton, N. R., Keyes, K. M., Krueger, R. F., Balsis S., Skodol, A. E., Markon, K. E., Grant, B. F., Hasin, D. S. (2012). An invariant dimensional liability model of gender differences in mental disorder prevalence: Evidence from a national sample. *Journal of Abnormal Psychology, 121*, 282 –288. doi: 10.1037/a0024780

Edens, J. F., Hart, S. D., Johnson, D. W., Johnson J. K., & Olver, M. E. (2000). Use of the Personality Assessment Inventory to assess psychopathy in offender populations. *Psychological Assessment, 12*, 132 –139. doi: 10.1037/1040-3590.12.2.132

Elrod, L. D., & Dale, M. D. (2008–2009). Paradigm shifts and pendulum swings in child custody: The interests of children in the balance. *Family Law Quarterly, 42*, 381 –418.

Emery, R. E., Otto, R. K., & O'Donohue, W. T. (2005). A critical assessment of child custody evaluations: Limited science and a flawed system. *Psychological Science in the Public Interest, 6*, 1 –29. doi: 10.1111/j.1529-1006.2005.00020.x

Erickson, S. K., Lilienfeld, S. O., & Vitacco, M. J. (2007). A critical examination of the suitability and limitations of psychological tests in Family Court. *Family Court Review, 45,* 157 –174. doi: 10.1111/j.1744-1617.2007.00136.x

Ezzo, F. R., Pinsoneault, T. B., & Evans, T. B. (2007). A comparison of MMPI-2 profiles between child maltreatment cases and two types of custody cases. *Journal of Forensic Psychology Practice, 7*, 29 –42. doi: 10.1300/J158v07n02_02

Ferrari, A. J., Baxter, K. J., & Whiteford, H. A. (2011). A systematic review of the global distribution and availability of prevalence data for bipolar disorder. *Journal of Affective Disorders, 134(1 –3)*, 1 –13. doi: 10.1016/j.jad.2010.11.007

Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment.* Washington DC: American Psychological Association. doi: 10.1037/10299-000

Grant, B. F., Hasin, D. S., Stinson, S. S., Dawson, D. A., Chou, S. P., Ruan, W. J., & Pickering, R. P. (2004). Prevalence, correlates, and disability of personality disorders in the United States: Results from the national epidemiologic survey on alcohol and related conditions. *Journal of Clinical Psychiatry, 65*, 948 –958. doi: 10.4088/JCP.v65n0711

Hetherington, E. M., Bridges, M., & Insabella, G. M. (1998). What matters? What does not?: Five perspectives on the association between the marital transitions and children's adjustment. *American Psychologist, 53*, 167–184. doi: 10.1037/0003-066X.53.2.167

Hynan, D. J. (2004). Unsupported gender differences on some personality disorder scales of the Millon Clinical Multiaxial-III. *Professional Psychology: Research and Practice, 35*, 105-110. doi: 10.1037/0735-7028.35.1.105

Jacobo, M. C., Blais, M. A., Baity, M. R., & Harley R. (2007). Concurrent validity of the Personality Assessment Inventory Borderline Scales in patients seeking dialectical behavior therapy. *Journal of Personality Assessment, 88*, 74 –80. doi: 10.1207/s15327752jpa8801_10

Langer, F. (2011). Using the NEO Personality Inventory in child custody evaluations: A practitioner's perspective. *Journal of Child Custody, 8*, 323-344.doi: 10.1080/15379418.2011.624440

Lansford, J. E. (2009). Parental divorce and children's adjustment. *Perspectives on Psychological Science, 4*, 140–152. doi: 10.1111/j.1745-6924.2009.01114.x

Longenecker, J., Genderson, J., Dickinson, D., Malley, J., Elvevag, B., Weinberger, D. R., & Gold, J. (2010). Where have all the women gone?: Participant gender in epidemiological and non-epidemiological research of schizophrenia. *Schizophrenia Research, 119*, 240–245. doi: 10.1016/j.schres.2010.03.023

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. doi: 10.1300/J158v01n02_02

Morey, L. C. (2007). *Personality Assessment Inventory: Professional manual (2nd Ed.)*. Lutz, FL: Psychological Assessment Resources.

National Conference of Commissioners on Uniform State Laws. (1979). *Uniform marriage and divorce act*. St. Paul, MN: West.

Posthuma, A. B., & Harper, J. F. (1998). Comparison of MMPI-2 responses of child custody and personal injury litigants. *Professional Psychology: Research and Practice, 29*, 437–443. doi: 10.1037/0735-7028.29.5.437

Rogers, R., Salekin, R. T., & Sewell, K. W. (1999). Validation of the Millon Clinical Multiaxial Inventory for Axis II disorders: Does it meet the Daubert standard? *Law and Human Behavior, 23*, 425–443. doi: 10.1023/A:1022360031491

Siegel, J. C. (1996). Traditional MMPI-2 validity indicators and initial presentation in custody evaluations. *American Journal of Forensic Psychology, 14*, 55–63.

Siegel, J. C., & Langford, J. S. (1998). MMPI-2 validity scales and suspected parental alienation syndrome. *American Journal of Forensic Psychology, 16*, 5–14.

Stricker, G. & Trierweiler, S. J. (1995). The local clinical scientist: A bridge between science and practice. *American Psychologist, 50*, 995–1002. doi: 10.1037/0003-066X.50.12.995

Strong, D. R., Greene, R. L., Hoppe, C., Johnston, T., & Olesen, N. (1999). Taxometric analysis of impression management and self-deception on the MMPI-2 in child-custody litigants. *Journal of Personality Assessment, 73*, 1–18. doi: 10.1207/S15327752JPA730101

World Health Organization, World Mental Health Survey Initiative (2007). *National comorbidity survey (NCS) and national comorbidity survey replication (NCS-R).* Retrieved from http://www.hcp.med.harvard.edu/ncs/

Received June 24, 2013, accepted August 4, 2013.