5-1-2005

"Faking good" response patterns on the MMPI-2 and the Child Abuse Potential Inventory

Stephen V. Flynn
Rowan University

"FAKING GOOD" RESPONSE PATTERNS ON THE MMPI-2 AND THE CHILD ABUSE POTENTIAL INVENTORY

by
Stephen Flynn

A Thesis

Submitted in partial fulfillment of the requirements of the
Masters of Arts Degree
of
The Graduate School
at
Rowan University
May 13, 2005

Approved by

Professor

Date Approved 5/1/05

ABSTRACT

Stephen V. Flynn
"Faking Good" Response Patterns on the MMPI-2 and the Child Abuse Potential Inventory
2004/05
Dr. Janet Cahill
Mental Health Counseling and Applied Psychology

A significant problem for forensic evaluations is the large number of clients who respond in a defensive manner. The present study examined defensive responding patterns on two self-report measures, the Minnesota Multiphasic Personality Inventory (MMPI-2) and the Child Abuse Potential Inventory (CAP). Subjects were child welfare clients undergoing a parental capacity evaluation. It was hypothesized that subjects would be less likely to attempt to portray themselves in an unrealistically positive manner (fake good) on the MMPI-2 than the CAP since the CAP items are more concerned with matters of child abuse and the MMPI-2 questions are primarily focused on Axis I and II diagnoses. It was further hypothesized that respondents who were court ordered would have higher faking good scores on both instruments. Results found the MMPI-2 had lower faking good scores compared with the CAP. No significant differences were found for any of the validity measures for court ordered status. Statistically significant patterns were found for court ordered status on some of the MMPI-2's clinical scales. Implications for the use of both instruments in forensic evaluations are discussed.

Acknowledgements

The author would like to thank all professors and fellow students who contributed to this study. The author is grateful to his Thesis Advisor Dr. Janet Cahill for her consistent guidance and collaboration during all phases of this research experiment and allowing the author to use the Rowan University Child and Family Assessment Clinic's Archival Data. Author would also like to thank Dr. Mary Louise Kerwin for her guidance and editing. A final thank you goes to fellow thesis advisee Shawn Carney for help during the data collection phase of this experiment.

Table of Contents

Chapter I: Literature Review

Psychologists are often asked to assist the courts and other legal venues with decisions regarding the resolution of child protection cases (Bathurst, Gottfried & Gottfried, 1997). A key component of these evaluations is an assessment of parental capacity. Assessing parental capacity can be viewed as having two aspects. The first is to determine the developmental and psychological needs of the child and the second is to assess the current ability of the parents or guardians to meet the child's needs (Heinze & Grisso, 1996).

According to Budd (2001), the major focus of parental capacity evaluations is obtaining information on what they understand, believe, know, do, and are capable of doing related to childrearing. Rather than completely focusing on the negative, there should also be a focus on what the parent's strengths are. Another major issue is utilizing a minimal parenting standard. This is defined as the minimum amount of care that is necessary to protect the well being of the child. For child welfare cases, this standard should be utilized instead of an optimal parenting standard, although this standard varies from state to state. There is a concern in regards to the lack of agreement as to what the minimum parenting standard should be.

The issue of a minimum parenting standard is of particular relevance to cases involving child abuse and neglect. Milner and Wimberley (1979) define child abuse as, "A situation in which a child is suffering from serious physical injury inflicted upon him by other than accidental means; is suffering harm by reason of neglect malnutrition, or

sexual abuse; is going without necessary and basic physical care; or is growing up under conditions which threaten his physical and emotional survival" (p.205). Budd (2001) describes risk factors for abuse and neglect as: psychiatric problems, teenage status, cognitive delays, substance abuse, criminal behavior, and chronic physical illness. Signs that children have suffered abuse or neglect include: non-organic failure to thrive, or lead intoxication, and unexplained injuries. The American Psychological Association (APA, 1994, as cited in Budd, 2001) recommends that when clinicians assess potential abusers, they focus on the current and potential abilities to function and how that correlates with meeting the needs of the child, the current relationship between the parent and child, the needs of the child psychologically and developmentally, clinical objectiveness and specific recommendation for intervention.

Budd (2001) further advocates the use of a multi-modal assessment protocol. Parental capacity evaluations are usually conducted in a coercive context that affects the reliability and validity of the results. To obtain the best objective assessment of the client, clinicians should apply multiple methods including clinical interviews, collateral information, observations of family interactions and standardized self-report measures.

While commonly used for forensic evaluations, self-report measures do have some limitations. Over reporting of symptoms (malingering or faking bad) and underreporting of symptoms (defensiveness or faking good) can result in invalid and misleading test results (Baer & Miller, 2002). Individuals are more likely to deliver response bias when tests are taken in settings that provide test takers with substantial incentives for distortions. These setting include child welfare cases where a loss of parental rights is a possible outcome (Baer & Miller, 2002). Clients undergoing an

evaluation in a child welfare context commonly underreport (and in some instances over report symptoms). Therefore, it is particularly important that assessment measures used in these settings have strong validity scales and indexes that can identify these response distortions.

One of the most frequently used self-report measures for child welfare evaluations is the Child Abuse Potential Inventory (CAP). The CAP was developed as a screening device for physical child-abuse (Milner & Wimberley, 1979). The CAP has a 77-item abuse scale organized in a forced choice agree/disagree format and has six subscales. The first subscale is distress, which has 36 items, and includes the assessment of affect including: anger, frustration, loss of self control, depression, and fear. The second subscale is rigidity, which has 14 items that assess a parent's beliefs in how much a child should always be neat, clean, quiet and obedient. The third subscale is unhappiness, which has 11 items that assess the parent's lack of personal fulfillment, pleasure, and amount of loneliness. The fourth subscale is problems with the child, which has six items that assess the thought that one's child is perceived as bad, slow, and a troublemaker. The fifth subscale is problems with family, which has four items that assess if the parent perceives other family members as having problems or if there is conflict among family members. The sixth subscale is problems with others, which has four items that assess a parent's belief that others make your life hard and cause pain. However, the only score that assesses the potential for child abuse is the overall abuse score.

The CAP was developed because of the need for an empirical approach in the identification of traits that distinguish child abusers (Milner & Wimberley, 1979). All of

the information that developed the CAP scales was taken from existing literature; in an attempt to base the scale on known theoretical constructs. The CAP abuse scale had an overall correct classification rate of 96% reported by Wimberley (1980), and there is a moderate, positive relation between elevated abuse scores on the CAP Inventory and later reports of abuse (Milner & Gold, 1984). The CAP attempts to measure parent pathology as well as interactional difficulties that are related to physical child abuse (Milner, 1986). A revised scale was developed (Milner, 1986) which added an18 item lie scale. A cross validation study determined that the use of the lie scale increased the overall correct classification rate (Milner, Gold, & Wimberley, 1986). Since the CAP assesses both parental pathology and interactional problems, substantial relationships should exist between the CAP and other measures of individual pathology and/or interpersonal problems such as the MMPI-2. Indeed, Milner, Charlesworth, Gold and Gold, (1988) found that the CAP was correlated with the MMPI depression scale and the Beck Depression Inventory.

The CAP has been empirically proven to differentiate physical child abusers from non-abusers. In a study by Milner, Gold and Wimberley (1986), 120 subjects were taken from an at-risk parent-child program and compared with 100 subjects from a non-abusive population. Results from an item analysis indicated that on the 77-item CAP abuse scale, 64 of the 77 items significantly differentiated abusers from control subjects. Further, discriminant analysis correctly classified 85.4% of the 220 subjects. Milner, Charlesworth, Gold, and Gold (1988) investigated the relationship between abuse potential and psychological variables measured by the Mental Health Inventory (MHI) (Gold et al, 1988). Results found significant correlations between the CAP abuse score

and each of the MHI scores ($p<.001$) in measuring the individual's ability to act out physically. More specifically, the CAP abuse scale had the strongest relationship with MHI summary and factor scales that measure psychological distress. The highest correlation between the CAP abuse scale and the five MHI factor scales was between abuse and the Loss of Behavioral/Emotional Control factor.

Holden, Foltz, and Willis (1989) looked at the efficacy of the CAP and the Parenting Stress Index (PSI) in distinguishing between abusing parents, neglecting parents, spouses of abusers, and parents who were referred for reasons other than confirmed abuse or neglect. They concluded that the CAP and the PSI appear to be more sensitive to abuse and high-risk situations than to neglect. Both the CAP total score and the PSI parent attachment scale scores were significantly lower for those parents who were reported for neglect. The study also examined gender differences. Overall, males reported lower levels of abuse potential (males, $M=133.8$, $SD= 93$; females, $M=161.5$, $SD=98.5$; $p<.05$). Males reported less parenting stress as well. PSI total stress score for men was $M=236.7$, $SD=45.2$; for females $M=252.6$, $SD=39.8$. This difference was significant ($p.<005$). At-risk parents for abuse reported higher levels of total parenting stress and specific parenting stress in relation to their sense of competence, relationship with spouse, health, and their child's mood than did the other participants. The results indicated that the PSI and CAP were assessing similar constructs, but that unique variance was also being assessed by each measure.

Blinn-Pike and Mingus (2000) used the CAP to explore whether children of adolescent mothers were at increased risk for child abuse compared to children of older mothers. The sample consisted of 105 adolescent mothers (mean age 17) who completed

the CAP abuse scale after the delivery of their babies. Results indicated the test-retest reliability was much lower for the adolescent mothers compared with the normative sample. There was moderate reliability in only two of the adolescent subscales. The authors concluded that additional research on the reliability and validity of this instrument for different populations was needed. Ornduff, Kelsey, Bursi, Alpert, and Bada (2002) conducted an investigation to determine the possible correlation between stressful life events and other risk factors in abuse and neglect. One outcome measured used was the CAP. Results indicated that higher CAP scores were related to exposure to violence in the family of origin and a lifetime use of street drugs. Taken together, these studies indicate that the CAP is a valid and reliable instrument for assessing the potential of physical abuse in most populations.

Another self-report instrument that is commonly used in parental capacity evaluations is the Minnesota Multiphasic Personality Inventory-2 (MMPI-2). The MMPI-2 has been reported as being the most frequently used measure in parenting and custody cases (Bagby, Buis, Fidler, Ncholson, & Radovanovic, 1999). According to Ackerman and Ackerman (1997), 31% of 201 doctoral level psychologists who responded to a questionnaire regarding child custody evaluations used the MMPI and the MMPI-2 when conducting screenings. A major strength of the MMPI-2 is the extensive validity scales included in the instrument. The MMPI was one of the first self-report inventories to include these important validity scales and an extensive amount of research has been conducted on their efficacy (Baer & Miller, 2002). Scales L and K and the F minus K index are most often used to detect the underreporting of symptoms (Baer, Wetter, Greene, Nichols, & Berry 1995). These scales have been found to be effective in

discriminating standard from underreported profiles (Baer, et al. 1995; Baer & Miller, 2002).

Additionally, new scales have been developed for the MMPI-2 in order to assess more subtle forms of underreporting symptoms. The socially desirable scale (Esd) includes items that 10 judges unanimously agreed on as a socially desirable response. The Test Taking Defensiveness scale (Tt) includes items for which judges agreed on the socially desirable response, but which only about half of a normative group endorsed in the socially desirable direction; O-S scale compares items that obviously detect psychopathology compared with more subtle ones (Baer & Miller, 2002). Some additional scales include: Other Deception (OD) scale, Positive Mental Health scale (PMH-4), Superlative (S) scale, Dissimulation scale (Ds-r), Malingering scale (Mp), and the Wiggins Social Desirability Scale (WSD; Baer et al, 1985; Bagby et al., 1995; Bagby et al, 1999). Research results have not been consistent on whether or not these instruments have been less or more effective than the traditional F and K scale. A study by Bagby et al. (1995) showed support for the continued use of the traditional validity scales for identifying faking good or faking bad profiles. This study did not find support for using the O-S-index. The results suggested that the traditional scales work as well or better than the Mp and Ds-r scales. Lastly, the F scale was found to be the strongest predictor in the detection of malingering. Bagby, Nicholson, Seeman, Rector, Rogers, and Buis (1997) concluded that the Esd, OD, and S scales were shown to be the best at predicting those patients who are trying to minimize their psychopathology. The L and K scales along with the F-K indexes were less successful in detecting fake good responses across a variety of situations. However, Bagby et al. (1999) found the two traditional

scales (F-K index) identified respondents who were faking good as well as some newer validity scales (WSD and S).

Although the MMPI-2 was specifically designed to measure personality pathology and predict psychiatric diagnosis (Quirk, Christiansen, Wagner, & McNutly 2003) and the CAP was designed to identify traits that distinguish child abusers (Milner & Wimberley, 1979), they are both commonly used in child protection settings. Since the problem of defensive responding is so common in these settings, the relative effectiveness of the two instruments in terms detecting a faking good response style is of interest. This study compared the validity scales on the MMPI-2 and the CAP lie scale to determine the percentage of clients who have a faking good profile. It is hypothesized that fewer individuals who take the MMPI-2 will have elevated validity scales compared with the lie scale on the CAP inventory. Individuals who fill out the MMPI-2 may feel the questions are less focused on their parenting ability and more focused on their personality. Thus, they may answer the questions more honestly and perhaps not realize the connection between their personal characteristics and abuse potential. On the other hand, individuals who take the CAP may notice that the questions are based around their parenting style and relationship with their child and may fake good on certain questions to produce an answer that sheds a more favorable appearance.

A second hypothesis focused on the court ordered status of the respondents. Since coerced clients have been found to have more defensive profiles, (Budd, 2000), it was hypothesized that court ordered clients would have high levels of faking good than referred clients.

Chapter II: Method

Test results were obtained from an archival data set in a university based clinic that services a child-welfare based population. All clients included in the data analysis were referred for parental capacity evaluations.

Subjects

Several subject pools were utilized. The first group completed the MMPI-2, the second group completed the CAP Inventory; and the third group completed both instruments. In both groups a total of ninety two subjects completed the MMPI-2. Females accounted for 70.6% of the subjects and males 28.2%. The age range was 19 to 73 and the mean age was 35.9. There was a relatively even split in whether or not subjects were court ordered to attend the clinic or were referred by their caseworkers. Forty eight subjects (52.2%) were court ordered to attend the clinic and 44 (47.8%) were not.

Sixty one subjects completed the CAP Inventory. Females accounted for 65.5% of the subjects and males 32.7%. Thirty-three of these respondents were court ordered (54.1%) and 28 were referred by their caseworkers (45.9%).

Forty one subjects completed both the CAP and the MMPI. Females represented 63.4% of this sample and 36.6% were male. The mean age for this group was 36.8. Demographic information for race, marital status, and income were not tracked in this data set.

Measures

A total of sixty one participants were given the CAP. A total of ninety two subjects completed the MMPI-2. The statistical program SPSS was used to calculate findings.

Procedure

Researchers gathered information regarding former adult clients of a university based clinic that services a child-welfare based population. Inclusion criteria included completion of the MMPI-2, CAP Inventory, or the completion of both. Clients were separated into three groups: group 1 included those participants who completed the MMPI-2; Group 2 included participants who completed the CAP Inventory; and Group three included those participants who completed both.

Chapter III: Results

The initial stage of this analysis was determined to see what percentage of this population had significantly elevated responses on validity scales of the MMPI-2.

L-Scale

The MMPI-2's Lie Scale (L) was used to determine how many subjects were faking good. A subject's response was considered invalid if they had an L-score equal to or greater than 80 (Hathaway et al, 2000). If the score is between 65 and 79 it is considered elevated and is likely to be invalid. Thirty six out of the 92 subjects (39.1%) who took the MMPI-2 had an L- Scale above 65. Of those, 12 had an L score of 80 or above (12.5% of the total sample).

TRIN-Scale

The MMPI-2's True Response Inconsistency Scale (TRIN), is used to see if the test taker is being inconsistent. If the inconsistency T score is above or equal to 80, the test is considered invalid. If the score is between 65 and 79 it is considered elevated and is likely to be invalid. Twenty-nine respondents (31.5%) had a TRIN score greater or equal to 65 and 5 (5%) had a score greater or equal to 80. There was one person who has scored above or equal to 80 on both the TRIN and L Scales. It was likely that this person either did not read the directions; was at a reading level below the normative group this test was constructed to fit; or responded to the questions in a random manner. Eleven subjects (10.1%) had both TRIN and L scales above 65. These test results are invalid, but it is difficult to determine if the problem is a deliberate attempt to fake good, or problems with understanding or completing the instruments. A total of 26 subjects (23.9%) had an

L scale above 65 and a TRIN scale below 65. These can be classified as the true faking good clients since their responses can not be accounted for by random responding. The results for these four scales were summarized in Figure 1.

Superlative- Scale

An elevated L scale on the MMPI-2 is generally interpreted as a relatively unsophisticated pattern of underreporting symptoms (Hathaway et al, 2000). The Superlative Scale (S) is a more subtle way of faking good resulting in an invalid response. There were 14 subjects (12.9%) who scored above or equal to 65. There was one subject who scored above 80. A total of 12 respondents (11.4%) had an S scale above 65 and a TRIN score below 65. These results are summarized in Figure 2.

Frequency-Scale

The MMPI-2 Infrequency scale was designed to identify clients who were faking bad, or had a hostile and negative view of the test. It can also be the result of poor reading abilities. In this sample, 17 subjects (15.6%) had an F scale greater or equal to 65.

Eleven respondents (10.1%) had both an L scale and an S scale above 65. All of these subjects had valid F and TRIN scores. Only four of the subjects also had an elevated L scale and none had an elevated S scale. However, 11 of the 17 also had an elevated TRIN score.

Figure I - Summary of the Standard Validity Scales for MMPI-2. All results are raw and out of a possible 92.

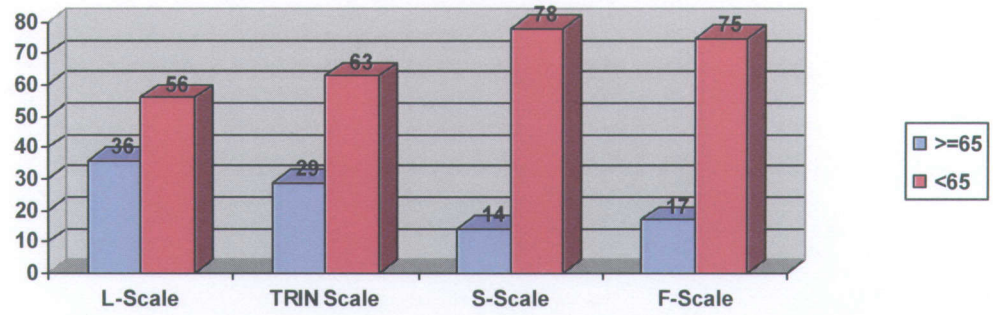
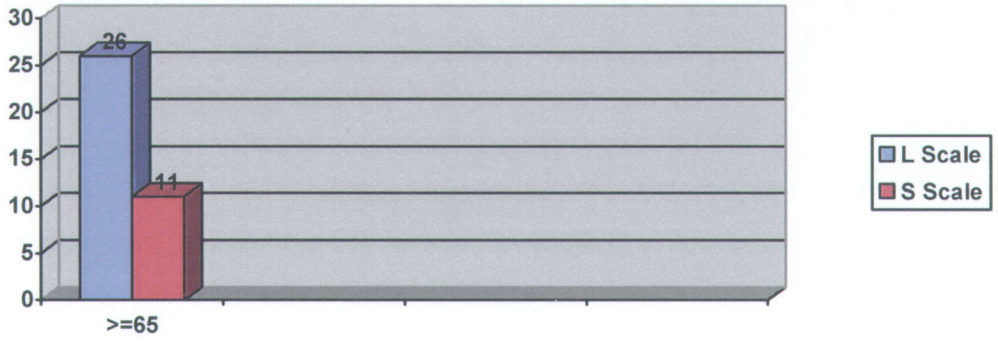


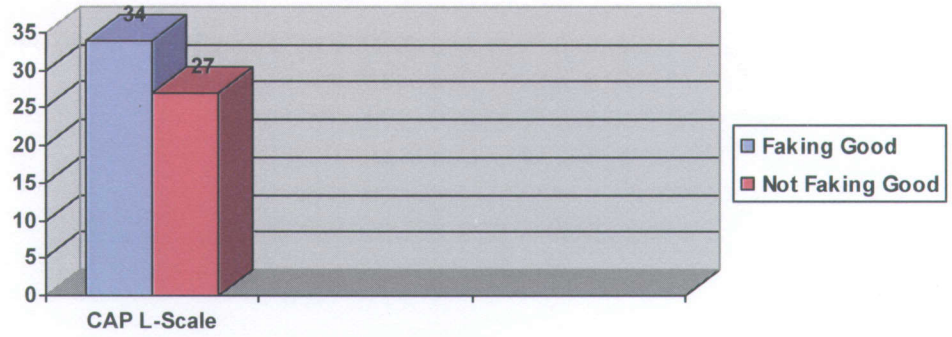
Figure 2 – Elevations in L and S scales with valid TRIN scores. All results are raw and out of a possible 61.



CAP Validity Scale

A client is classified as faking good on the CAP if their Lie scale is equal to or greater than 7 (Milner & Wimberely, 1979). Out of the 61 total subjects who completed the CAP, 34 subjects (55.7%) had an L scale above or equal to 7 in the subject pool. These results were summarized in Figure 3.

Figure 3 – Comparison of faking good and non-faking good subjects on the CAP Inventory. All results are raw and out of 61 possible.



In order to control for possible demographic or clinical differences between the single test subject pools, the subjects who took both tests were examined. Forty-one respondents completed both the MMPI-2 and the CAP. Of those, 22 (53.7%) had an elevated Lie scale on the CAP. Fifteen had an elevated L scale on the MMPI >= 65 and TRIN score below 65 (36.6%). Six subjects had an S scale of >= 65 and a TRIN score below 65 (14.6%). Six had an F scale >=65 (14.6%).

In terms of the consistency of faking good between the two measures, 15 subjects had both an elevated Lie scale on the CAP and an L scale on the MMPI. No subject had an elevated L scale, and a Lie scale below seven. Subjects who took both tests and had an elevated Lie Scale on the MMPI-2 also had an elevated L scale on the CAP. However, seven subjects (1.7%) had a CAP Lie scale above seven and a Lie scale below 65. In terms of the S scale, seven subjects (1.7%) had a CAP Lie scale above seven and an S scale below 65. Only one subject had an S scale above 65 and a CAP Lie scale below seven.

Court Ordered Status and Validity Scales

In order to test for differences between the court ordered and non-court ordered groups in terms of their responses to the validity scales, independent sample t-tests were run which compared court ordered status and the CAP Lie Scale and the TRIN, S and L scale. No statistically significant differences were found. (Lie t= -1.65, df= 39; p<.105; L scale t= -.719, df= 39, p<. 477; S scale t= -.500, df= 39, p<620; TRIN t=.509, df=39, p<.614).

Clinical scales on the MMPI-2 were analyzed to determine whether or not court ordered clients presented with more pathology than non-court ordered subjects. Scoring above or equal to 65 on all scales determined clinical significance. Twelve court-ordered subjects presented with an elevated Paranoia Scales compared to six non court ordered subjects. Five court ordered subjects presented with an elevated Depression Scale compared to six non court ordered subjects. Seven court-ordered subjects presented with an elevated Anger Scale compared to one non court ordered subjects. There were 13 court ordered subjects presenting with an elevated Cynicism Scale compared to eight non court ordered subjects. Four court ordered subjects presented with an elevated Anxiety Scale compared to three non court ordered subjects.

A Chi-Square was used to examine potential gender differences between court ordered and non-court ordered clients. Results were not significant χ^2=2.50, df=2, p<.286.

Court Ordered Status and MMPI-2 Clinical Scales

The final analysis examined whether or not there were differences in court ordered status on the clinical scales of the MMPI-2. Independent t tests on the MMPI-2

subjects revealed no significant differences except for the Masculine Feminine (MF) scale Court ordered clients had a significantly higher elevation on this scale ($t=2.76$ $df=89$; $p<.007$).

Chapter IV: Discussion

The present study provides good support for the effectiveness of the CAP and MMPI-2 validity scales, within a child and family assessment based population. Both of these inventories were successful in identifying a large number of clients who were trying to underreport their symptoms. Due to the high levels of faking good that was reported, it may be appropriate for clinical evaluators to put more weight behind the clinical interview and direct observations of behavior than self-report assessment tools.

The CAP had higher levels of faking good compared with the MMPI-2. Over half (55%) of the subjects who took the CAP Inventory showed an elevated Lie Scale compared with 39.1% on the MMPI-2. There may be several explanations for this finding. It may be possible that the CAP is more directly related to child abuse and neglect and therefore more threatening to the client. A second explanation could be that the CAP may have a more sensitive validity scale compared with the MMPI-2. Another important question is which test was more efficient at picking up invalid responses. After the subjects who took both tests were examined it was clear that the CAP lie scale was picking up a larger amount of subjects who were faking good compared with the MMPI-2 L-Scale. In fact, the CAP identified 22 subjects (53.7%) with an elevated lie scale compared with the 15 subjects (36.6%) picked up by the MMPI-2 L-Scale. On the other hand, the MMPI-2 F and S scales were able to detect an additional 12 subjects who presented with defensive profiles. The CAP may have a more highly sensitive L-scale, but it does not detect the range of faking good subtypes picked up by the L, TRIN, S and F scales on the MMPI-2. It can be argued that

the MMPI-2 actually detected more faking good clients overall because of the range of validity scales it has.

Although, it is evident that clients who take the CAP and MMPI-2 in this coercive setting tend to paint an unrealistically positive picture of themselves, there are those clients who do answer honestly. The results for these clients are valid and the test results give evaluators useful information regarding their abusive behavior and/or pathology.

Both of these tests measure different constructs. The CAP measures abuse and neglect potential while the MMPI-2 measures psychopathology. The CAP informs evaluators the extent of potential physical abuse while the MMPI-2 may give evaluators clues as to the potential causes of this elevated abuse potential. These tests together will give evaluators good insight into the abuse potential and possible reasons as to why the abuse is taking place. Since they appear to provide different types of information the use of both tests in the same protocol appears to provide incremental value to the overall assessment.

In detecting fake good candidates on the MMPI-2, the L-scale was shown to be the most valuable validity scale. Not only did the L-scale pick up the most invalid profiles, but it is also the only validity scale specifically designed to detect clients who fake good. However, it is possible that elevations in the L scale are the results of other factors besides a deliberate attempt by the client to present him or herself in a very positive light. It may also be that the client is responding in a random or inconsistent manner. That is why the TRIN scale is also extremely important in the process of detecting clients whom fake good. If a TRIN score is greater than or equal to 65, it is highly likely that the client who took the test answered the questions in a random

response style, and thus would eliminate them from the fake good group. Thus, the two validity scales mentioned are the most valuable for detecting clients with fake good profiles.

Overall, the pattern of faking good for this population is quite large. This is not surprising given the coercive nature of the evaluations. Clients seem to fake good more often and to a greater extent to questions directly related to their relationship with their family and abuse/neglect potential compared with questions based on psychopathology. Clients in this population may not see the relationship between certain affect (depression and anxiety) and likelihood for abuse.

There were no significant differences found for the validity measures for court ordered status. Statistically significant patterns were found for court ordered status on only one of the MMPI-2 clinical scales. The absence of differences in faking good levels between court ordered and non –court ordered clients could be because clients view the process as coercive even when they are referred by their caseworkers. This perspective has some validity since failing to comply with the service recommendations of a child welfare agency could also have led to the loss of access to their children.

There was a statistical difference on the masculine-feminine (MF) scale of the MMPI-2. This could be the result of statistical error since a number of t-tests were conducted. The MF Scale was originally developed by Hathaway et, al (2000) to identify males with purported homosexual traits. In the original version of the test, the authors identified only a very small number of items that differentiated homosexual from heterosexual males (Hathaway et al, 2000). Subsequent research has found that scores on this scale are also related to intelligence, education, and socioeconomic status (Hathaway

et al, 2000). For example, it is not uncommon for male college students and other college-educated males to obtain T-scores in the 60 to 65 range. Scores that are markedly higher than expected for males, based on the persons' intelligence, education, and social class, should suggest the possibility of sexual concerns and problems. High scores are very uncommon among females. When they are encountered, they generally indicate rejection of the traditional female role. Of the 60 items in the original scale 5, 56 have been maintained in the MMPI-2. It is possible that the females whom are being court ordered for evaluations are more likely to have rejected the traditional female role and have stopped any nurturing and empathic behavior and instead turned to child neglect and abuse. However, there is insufficient research to make strong conclusions about the meaning of this difference with this population.

Overall, the clinical profiles of court ordered and non-court ordered clients were quite similar. This seems to indicate that court orders do not necessarily indicate a more dysfunctional group. This may mean those who are court ordered are no more pathological or likely to abuse as those who are not court ordered.

In conclusion, the present study provides good support for the effectiveness of the CAP and MMPI-2 validity scales within a child and family assessment based population. As predicted, the coercive atmosphere in the population allotted for a significant amount of faking good on both inventories. The results confirmed the hypothesis that there would be more faking good on the CAP inventory compared with that of the MMPI-II. The lack of variation between court ordered and non-court ordered clients was unexpected. This seems to indicate that court orders do not necessarily indicate a more dysfunctional group. It is probable that the complexity of issues in child custody cases

will continue to increase. Hopefully, additional research on the MMPI-II and the CAP is undertaken to build on the results of this study.

Chapter V: References

Ackerman, J.M., & Ackerman, C. M. (1997). Custody evaluation practices: a survey of experienced professionals (revisited). *Professional Psychology: Research and Practice, 28,* 137-145.

Baer, R.A., & Miller, J. (2002). Underreporting of psychopathology on the MMPI-2: a meta-analytic review. *Psychological Assessment. 14, (1),* 16-26.

Baer, A.R., Wetter, W.M., Greene, R., Nichols, S.D., & Berry, T.D. (1995). Sensitivity of MMPI-2 validity scales to underreporting of symptoms. *Psychological Assessments, 7,* 419-423.

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

Bagby, M.R., Buis, T., & Nicholson, A. R. (1995). Relative effectiveness of the standard validity scales in detecting fake-bad and fake-good responding: replication and extension. *Psychological Assessment, 7,* 84-92.

Bathurst, K., Gottfried, A.W., & Gottfried, A. E. (1997). Normative data for the MMPI-2 in child custody litigation. *Psychological Assessment, 9, (3),* 205-211.

Budd, K. S. (2001). Assessing parenting competence in child protection cases: a clinical practice model. *Clinical Child and Family Psychology Review, 4, (1),* 1-17.

Hathaway, S. R., McKinley, J. C., Meehl, P. E., Drake, L. E., Welsh, G. S., & MacAndrew, C. (2000). Construction of the original MMPI. J. N. Butcher (Ed.), Basic sources on the MMPI-2. (pp. 1-100).

Heinze, M.C., & Grisso, T. (1996). Review of instruments assessing parenting competencies used in child custody evaluations. *Behavioral Sciences and the Law, 14,* 293-313.

Holden, E.W., Foltz, L., & Willis, J.D. (1989). Child abuse potential and parenting stress: relationships in maltreating parents. *A Journal of Consulting and Clinical Psychology, 1, (1),* 64-67.

Milner, J. S. (1991). Medical conditions and child abuse potential inventory Specificity. *A Journal of Counseling and Clinical Psychology, 3 (2),* 208-212.

Milner, J. S., Charlesworth, R. J., Gold, R.G., Gold., & Gold R.S.(1988). Convergent validity of the child abuse potential inventory. *Journal of Clinical Psychology, 44, (2),* 281-285.

Milner, J.S., & Gold, G.R. (1984). Predictive validity of the child abuse potential inventory. *Journal of Counseling and Clinical Psychology. 52, (5),* 879-884.

Milner, M. S., Gold, R.G., & Wimberley, C.R. (1986). Prediction and explanation of child abuse: cross-validation of the child abuse potential inventory. *Journal of Counseling and Clinical Psychology, 54, (6),* 865-866.

Milner, S.J., & Wimberley, R.C., (1979). An inventory for the identification of child abusers. *Journal of Clinical Psychology, 35, (1),* 95-100.

Pike, L. B., & Mingus, S. (2000). The internal consistency of the child abuse potential inventory with adolescent mothers. *Journal of Adolescence, 23,* 107-111.