Differentiating Faking-good Parents from Non-faking Parents on the Child Abuse Potential Inventory

Amanda H. Costello

West Virginia University

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Recommended Citation
https://researchrepository.wvu.edu/etd/3343
Differentiating Faking-good Parents from Non-faking Parents on the Child Abuse Potential Inventory

Amanda H. Costello, B.S.

Thesis submitted to the Eberly College of Arts and Sciences at West Virginia University in partial fulfillment of the requirements for the degree of Master of Science in Psychology

Cheryl B. McNeil, Ph.D., Chair
William J. Fremouw, Ph.D.
Amy L. Gentzler, Ph.D.

Department of Psychology
Morgantown, West Virginia 2011

Keywords: Child abuse and neglect; parental fitness evaluations; faking good behavior
ABSTRACT

Differentiating Faking-good Parents from Non-faking Parents on the Child Abuse Potential Inventory

Amanda H. Costello

Parental fitness evaluations remain common during child abuse investigations, as evidenced by the 19% of victims of maltreatment receiving an evaluation in 2007 in the United States (U.S. Department of Health and Human Services, 2009). A “best practice” approach to parental fitness evaluations consists of using a comprehensive assessment including clinical interviews, risk assessment measures, and behavioral observation (Budd, 2001; 2005). The Child Abuse Potential Inventory (CAP; Milner, 1986) is one risk assessment measure that is often used during evaluations. Unfortunately, abuse risk at pre- and post-treatment as measured by the CAP may be confounded by faking-good responding from the parent. Currently, there is a small amount of literature examining the characteristics of parents who fake good on the CAP, and the implications of invalid profiles at pretreatment. The present study examined differences between parents with a faking or non-faking profile on the CAP across multiple pretreatment variables including demographic information, psychopathology, and behavioral observation. Parents differed significantly on level of IQ and depression scores, but no significant differences were found on any other variable. Additionally, recidivism rates were examined for both groups of parents at posttreatment. Nine (25%) parents with an elevated Faking-good index and 16 (22.7%) parents without an elevation recidivated at posttreatment, and no significant differences were found between groups. Implications of study outcomes, and the emphasis for a multimethod approach to parental fitness evaluations will be discussed.
TABLE OF CONTENTS

I. Introduction (pp. 1-2)
   a. Predicting Future Abuse in Fitness Evaluations (pp. 2-4)
   b. Socially Desirable Responding and its Consequences (pp. 4-7)
   c. Characteristics of Faking-good Parents (pp. 7-20)
   d. Treatment Outcomes of Faking-good Parents (pp. 21)

II. Statement of the Problem (pp. 21-23)

III. Exploratory Questions (pp. 23-24)

IV. Method
   a. Participants (pp. 24-26)
   b. Measures (pp. 26-30)

V. Analyses (pp. 30-31)

VI. Results (pp. 31-35)

VII. Discussion (pp. 35-40)
   a. Limitations and Future Directions (pp. 40-42)

VIII. References (pp. 43-51)

IX. Table 1: A Comparison of Faking-good versus Non-faking parents Across Demographic Information, Psychopathology, Child Welfare System Involvement, and Behavioral Observation at Pretreatment (pp. 52)

X. Table 2: Comparing Recidivism Rates and Behavioral Observation Data across Faking and Non-faking Parents at Posttreatment (pp. 53)

XI. Appendix A (pp. 54)

XII. Author Note (pp. 55)
Introduction

The assessment of child abuse potential in maltreating parents can be a highly stressful and difficult situation for both the offending parent and the mental health professional. Parents who are involved in the Child Protective Services system are often court-ordered to participate in a parental fitness evaluation, which, by best practice, is a comprehensive procedure utilizing a range of measures including clinical interviews, self-report assessment, and behavioral observation (Budd, 2001; 2005). The Child Abuse Potential Inventory (CAP; Milner, 1986) is a self-report abuse risk assessment commonly used in parental fitness evaluations (Budd, Heilman, & Kane, 2000; Haskett et al., 1995). Because these evaluations can result in serious consequences for the offending parent (e.g., termination of parental rights), many individuals engage in faking-good responding on risk assessment measures. The CAP has remained a strong risk assessment because it contains validity scales to identify faking-good responding. An invalid CAP due to faking-good can complicate parental fitness evaluations because it creates difficulty for mental health professionals in understanding a parent’s “true” behavior when making predictions for the future welfare of the child.

Although invalid CAPs due to faking-good add a multitude of complications to parental fitness evaluations, there is limited empirical information describing characteristics of parents who engage in this behavior. This differs from a strong research base for similar measures such as the Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Butcher, Graham, Ben-Porath, Tellegen, Dahlstrom, & Kaemmer, 2001). Because the CAP is widely used in parental fitness evaluations, more faking research is needed using this measure. Additionally, a limitation of the majority of studies on faking behavior is the omission of a behavioral observation component. Many research studies utilizing both the CAP and the MMPI-2 have relied solely on parental
self-report, even though authors of these studies have emphasized the use of a multimethod approach (e.g., Cooke, 2010; Medoff, 1999). This study attempts to examine the differences between faking and non-faking parents using the CAP validity scales on multiple variables including demographic information, self-report measures of psychopathology, and behavioral observation. It also aims to examine differences in treatment outcomes between parents who fake and those who answer honestly on the CAP.

Predicting Future Abuse in Fitness Evaluations

The serious and potentially life-threatening consequences of child abuse and neglect have demonstrated the importance of predicting future abuse in assessment and treatment (Milner, 1991). Abuse potential assessment is often conducted in the form of a parental fitness evaluation in either the home or a clinical setting. Parental fitness evaluations were established to determine the minimal standards of parenting for abusive and neglectful parents, and are often composed of multiple assessment measures including a clinical interview, behavioral observation, and parental self-report measures (Azar, Lauretti, & Loding, 1998; Barone, Weitz, & Witt, 2005; Budd, 2001; 2005; Carr, Moretti, & Cue, 2005; Rohrbagh, 2008). Court-ordered investigations are common, as evidenced by the 19.0% of victims of maltreatment who received evaluations in 2007 in the United States (U.S. Department of Health and Human Services, 2009). Implications of parental fitness evaluations vary, and can result in determination of parental rights, referral of families for treatment, settling child custody issues, and ensuring the child’s protection (Haskett et al., 1995). In some cases, parental rights are terminated (Azar et al., 1998; Barone et al., 2005). Thus, these evaluations potentially have lifelong and life-altering outcomes for both the parent and the child. Yet these evaluations are fallible in that they require forecasts about future behavior that is not necessarily predictable (Barone et al., 2005).
Despite the serious potential outcomes from parental fitness evaluations, researchers are still beginning to establish a uniform standard of care for addressing parental fitness. To further add to the complexity of evaluations, setting variables (e.g., location such as the home versus clinic), demand characteristics, parental characteristics (e.g., socially desirable behavior), and even assessor bias are all important variables to consider when predicting the future welfare of the child (Budd & Holdsworth, 1996; Cooke, 2010; Kazdin, 2003; Medoff, 1999; Milner, Murphy, Valle, & Tolliver, 1998; Wolfe, 1988). Additional research for determining the necessary components of parental fitness evaluations has been emphasized, and researchers are beginning to identify these components (Budd, 2001; 2005; Carr et al., 2005). One such component is risk assessments (e.g., Child Abuse Potential Inventory, Milner, 1986; Parenting Stress Index, Abidin, 1995) for the identification of future child abuse. Risk assessments have been cited as being commonly used in parental fitness evaluations and have utility in predicting future abuse (Chaffin & Valle, 2003). Furthermore, they do not necessarily hurt client-therapist rapport, as compared to a clinical interview with the parent, because they do not require mandatory abuse reporting. They also do not require the need for chart review/child welfare database searching, and because they are able to predict future risk, do not necessarily require longitudinal studies (Chaffin & Valle, 2003).

One of the most commonly used risk assessments is the Child Abuse Potential Inventory (Milner, 1986) (Budd, Heilman, & Kane, 2000; Haskett et al., 1995). The CAP is a 160-item self-report measure in which individuals are instructed to either agree or disagree with each item. Each item conforms to a third-grade reading level (Milner, 1986). The CAP contains six abuse risk domains including distress, unhappiness, rigidity, problems with child/self, and problems from others (Milner, 1986). It also yields a broadband abuse risk score, which makes it a
comprehensive measure for future abuse incidence rates. The measure contains specific instructions to answer honestly, and its manual recommends that experimenters emphasize these instructions before beginning test administration (Milner, 1986). The CAP has remained a sound measure by demonstrating strong construct validity (Haskett, et al., 1995) and has been correlated with measures associated with multiple parental psychosocial difficulties (e.g., a history of maltreatment, distress and low self-esteem, and dysfunctional parent-child interactions; see Aragona, 1983; Fulton, Murphy, & Anderson, 1991; Haskett et al., 1995; Miller, Handal, Gilner, & Cross, 1991; Milner, Charlesworth, Gold, Gold, & Friesen, 1988; Stringer & LaGreca, 1985). A prospective study of the CAP identified future abuse risk for parents scoring above the clinical cutoff on the physical abuse scale (Chaffin & Valle, 2003; Milner, 1994). Additionally, parental self-reported physical abuse risk has found to be associated with negative parent-child interactions and an increased authoritarian parenting style (Haskett et al., 1995), leading to the usefulness of this measure in predicting negative or risky parenting styles.

Socially Desirable Responding and its Consequences

Although there have been well-established assessments for the prediction of future abuse incidence rates, parental fitness evaluations may still be confounded by socially desirable responding, or a parent engaging in faking-good behavior (Carr et al., 2005). Faking-good behavior can be described as parents engaging in an attempt to “minimize their personal and family problems” (Milner & Crouch, 1997, pp. 633) and is often used to derive some sort of benefit. Parents may engage in faking-good behavior to decrease the chances of having their parental rights temporarily or permanently terminated. Parents may also want to appear socially desirable to mental health professionals (Bennett et al., 2006; LaFiosca & Loyd, 1986; Miller-Perrin & Perrin, 1999) leading to invalid reporting or changes to their typical parenting behavior.
Many clinicians report that faking-good behavior is one of the most difficult areas to address in forensic psychology (Otto, Edens, & Barcus, 2000). In the realm of research these individuals are often described as “apprehensive subjects” (Kazdin, 2003) who feel pressure to distort their behavior to impress researchers. In applied settings, parents involved in child abuse allegations also are at risk for demonstrating “apprehensive subject” characteristics.

The CAP was developed to include three validity scales used to address distorted responding. The measure includes a Random Responding, Inconsistency, and Lie scale. The Random Responding scale was developed to identify respondents who were either confused by items or answered without considering item content (Milner, 1986). The Inconsistency scale was developed in conjunction with the Random Responding scale to produce a more accurate invalid profile (e.g., reducing Type I error by classifying non-random respondents as random respondents; see Milner, 1986). The Lie scale is the scale most closely associated with socially desirable responding (Milner, 1982), and becomes an important piece of information during parental fitness evaluations. Milner and colleagues conducted multiple studies to develop the Lie scale on the CAP. Early studies administered items indicative of socially desirable responding to undergraduate students in an attempt to develop an appropriate scale. One study tested undergraduate participants using a CAP containing 47 socially desirable items (e.g., “I never lose my temper”; “I always tell the truth”). Of the 47 original items, 30 items were chosen (Milner, 1982; Milner, 1986). In a follow-up study, the 30 items were given to participants from a parent-teacher association in North Carolina, and the original set was pared down to 18 items. Some examples of these items included “I never raise my voice in anger” and “I never lose my temper” (Milner, 1982; Milner, 1986).
The 18 items were then integrated into the full scale (e.g., 160-item CAP) and administered to a normal adult cohort (e.g., without a history of maltreatment). Participants were required to answer each item using a forced-choice “yes/no” scale, and socially desirable items were assigned a score of one. Therefore, a higher score on the Lie scale was associated with socially desirable responding (Milner, 1986). For parents with a high school diploma or above, a clinical cutoff of 7 indicating an invalid profile was developed (Milner). For parents with less than a twelfth grade education, a clinical cutoff of 8 was used to account for potential differences in reading and comprehension.

Construct validity was assessed in a study conducted by Robinson and Milner (1983) using undergraduate student participants. The authors compared scores on the CAP Lie scale with scores on a modified version of the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1943) and the Barron ego-strength scale (Barron, 1953) to examine the ability of each scale in correctly identifying socially desirable responding. The CAP was able to correctly distinguish between 84-95% of participants with an elevated Lie scale. The authors were able to replicate results utilizing a similar procedure, recruiting parent participants in their follow-up study (Robertson & Milner, 1985).

The CAP also includes a specific Faking-good index, which consists of an elevated Lie scale (e.g., a score of 7 or above) and a Random Responding scale within normal limits. Typically about 7-11% of general population parents have an elevated Faking-good index (Milner, 1986), but this behavior may change depending on the clinical or forensic setting. Researchers conducted additional studies to assess the validity of faking-good behavior utilizing the CAP Lie scale. Matthews (1984) demonstrated the association between Lie scale and the L scale (i.e., lie scale) on the MMPI using parent participants with a mildly abusive background
(although “mildly abusive” was not defined). The lie scales on the CAP and MMPI were found to be significantly positively associated (.48), and the CAP was able to be distinguished from the F scale (measuring confusion and self-deprecation) and the K scale (measuring defensiveness) (Matthews, 1984). The studies conducted with the MMPI were especially important to include, as the MMPI has been widely used in parental fitness evaluations (Carr et al., 2005). Finally, the CAP lie scale was tested against the lie scale on the Eysenck Personality Inventory (EPI; Eysenck & Eysenck, 1968) and a significant positive correlation (.51) was demonstrated between the two scales (see Pruitt, 1983).

**Characteristics of Faking-good Parents**

Although faking-good behavior is a legitimate concern, and the CAP has a strong empirical basis for its Lie scale, there is limited research utilizing the CAP validity scales when examining behaviors which distinguish faking parents from non-faking parents. This is especially critical information given the sensitive nature of parental fitness evaluations and the difficulty experienced by mental health professionals in determining the rights of parents involved in the child welfare system. Mental health workers using the CAP who are unable to distinguish between faking and non-faking parents may be at risk for making decisions that are inappropriate for family placement, and may place a child in danger for future abuse.

Therefore, it is important to describe studies which have begun to address faking-good behavior using the CAP and other risk assessment measures or personality inventories. Many studies using personality inventories have focused primarily on prisoners or inpatient psychiatric patients, but there are studies that have been conducted on parents who are involved in either Child Protective Services (CPS) or in child custody evaluations. Some studies have focused specifically on instructional variables, while others have examined individual parental
characteristics in relation to socially desirable responding. Instructional variables study the effect of variations in the presentation of study procedures by experimenters on participant behavior (see Kazdin, 2003). Studies on individual parental characteristics have focused on information derived from self-report measures. A study conducted by Milner and Crouch (1997) examined the influence of instructional cues in faking-good responding on the CAP by altering prompts presented to parent participants. When cued to “fake-good” by an experimenter on the CAP, both general population parents and parents who were at-risk for physical abuse were able to distort their profiles in a socially desirable manner (Milner & Crouch, 1997). This differed when the experimenter cued parents to answer honestly. Interestingly, parents in both groups were unable to produce differences in specific scales (e.g., Rigidity and Problems with Child/Self). This study showed that altering experimenter behavior can result in changes in responding. Participants in this study demonstrated sensitivity to instructional cues which may be present (albeit in a more subtle manner) in parental fitness evaluations. But, it did not address the issue of differences between the two groups.

Differences between faking and non-faking parents have been studied in more detail using individual parental characteristics (e.g., intelligence, parental distress, attitudes toward children). Unfortunately many of these studies have used personality inventories like the MMPI-2 over risk assessment measures like the CAP. Budd and colleagues (2000) did conduct a study using the CAP and examined psychosocial correlates of abuse risk in 75 adolescent mothers. In addition to the CAP, the authors used multiple measures associated with child abuse risk. These included the Wide Range Achievement Test-Revised (WRAT-R; Jastak & Wilkinson, 1984) to test for reading ability, the Parent Opinion Questionnaire (POQ; Azar, Robinson, Hekimian, & Twentyman, 1984) to assess for unrealistic parenting expectancies, the Arizona Social Support
Interview Schedule (ASSIS; Barrera, 1981) to assess for an individual’s exposure to social contacts across a range of settings, the Symptom Checklist 90-Revised (SCL-90-R; Derogatis, 1983) to assess for parent psychopathology, and Home Observation for the Measurement of the Environment (HOME Inventory; Caldwell & Bradley, 1984) as a measure of behavioral observation of parent-child interactions and the quality of the child’s well-being in the home. Adolescent mothers in the study were wards held by the Illinois Department of Children and Family Services (DCFS; Budd et al., 2000) recruited by DCFS caseworkers or individual therapists in private practice. These adolescent mothers retained their parental rights, such that their children continued to live in their care. Any serious risk factors observed in the home (e.g., severe psychopathology or the witnessing of domestic violence) were reported by researchers; otherwise, mothers did not have any direct threats to their parenting rights.

Despite the fact that mothers did not experience any direct threats to their parenting rights, the authors found that 19% of mothers in the study yielded an invalid profile on the CAP (Budd et al., 2000). The high percentage of invalid profiles was obtained even after the authors followed standard protocol developed by Milner and colleagues (1986) to include a lie score of 8 (instead of 7) for parents with a twelfth grade education or less. All of the invalid profiles had an elevated Faking-good index (Budd et al., 2000). Group comparisons between mothers with an elevated abuse risk score, scores within normal limits (e.g., a valid profile without an elevated abuse risk score), and invalid profiles were conducted on multiple variables including years of education, and scores on the WRAT-R, ASSIS, and SCL-90-R (Barrera, 1981; Derogatis, 1983; Jastak & Wilkinson, 1984). Mothers with invalid profiles only differed from their counterparts in regard to reading achievement (Jastak & Wilkinson, 1984). Specifically, mothers with invalid
profiles had a lower WRAT-R score than both the mothers in the elevated and normal groups. Interestingly, there were no differences in years of education completed.

Outcomes from the Budd and colleagues (2000) yielded some interesting findings. The adolescent mothers in the study with invalid profiles may not have been able to understand the measures administered, leading to proneness for socially desirable responding. This could be concerning given the tendency for faking individuals to be portrayed as intelligent, manipulative parents who are savvy at navigating the child welfare system. Additionally, past research has cited that many individuals involved in the forensic setting possess low level of education and intelligence (see Klinge & Dorsey, 1993). But, parents in this study were administered each assessment orally, which should have accounted for differences in reading ability. The researchers also correctly followed the adaptations to the lie scale cutoff for mothers with a twelfth grade education or less. Because there was a disconnect between scores on the WRAT-R and the assessment administration and scoring, it yields confusing findings as to the importance of reading ability for these mothers in understanding the items on the CAP.

The adolescent mothers also did not experience a direct threat to termination of their parental rights, and were not participating in a treatment outcome study that could have future parenting implications. As there was not a risk for termination, mothers in the study may not have had the motivation to fake-good on the CAP, making these mothers very different from other “at-risk” or abusive parenting groups in research or clinical settings. It is notable to observe that despite not having an open CPS case, 19% of the sample still scored in the fake-good range. This could have been due to participation in a research study and mere contact with an experimenter in the mental health field. The sample studied remains important; faking-good responding may increase as a function of setting or characteristics of participants (e.g., parents
who do not have any direct threats to their parenting rights versus parents who may lose their children), so it is important to understand sample characteristics when assessing faking-good.

One important variable missing from the comparisons between the faking and non-faking parents in the Budd and colleagues (2000) study was the behavioral observation data. The HOME assessment measure used by Budd and colleagues contains six subscales (e.g., responsivity, acceptance, organization, learning materials, involvement, and variety; Caldwell & Bradley, 1984) which may have provided much more information on differences in parenting styles across the three groups. Unfortunately the authors did not analyze and report the relationship between behavioral categories measured by the HOME scale and faking-good responding on the CAP. The authors did note that an important limitation of their study was the fact that most of the data were collected as adolescent self-report (Budd et al.). Given that behavioral observation often can yield a less biased, and more realistic snapshot of behavior, it should have been included as a means of comprehensively assessing differences in parenting behavior. The adolescent mothers in the study may have been able to present themselves in a favorable manner during self-report assessment, but may not have had the skills to maintain behavior under a different context. Although the Budd and colleagues study strengthened the literature for psychosocial correlates of abuse risk, it also left many “future directions” in faking research, notably the reading and comprehension ability of parents, the setting of the research, and potential differences in self-report and behavioral observation data.

An example of a study utilizing behavioral observation was conducted by Bennett, Sullivan, and Lewis (2006) examining faking behavior of mothers involved in CPS. In this study, the association between mothers’ self-report responding on the Parent-Child Conflict Tactics Scale (CTSPC; Straus, Hamby, Finkelhor, Moore, & Runyan, 1998) and their behavioral
observations was addressed. Although nearly 30% of abusive mothers did deny involvement in CPS, behavioral observation and self-report measures were unable to distinguish between those mothers and control mothers (i.e., those who did not deny involvement). The CTSPC did show some differences (e.g., through the neglect, physical assault, and psychological aggression subscales) between mothers who admitted prior CPS history and those who did not, but denying mothers were still unable to be distinguished. The Bennett and colleagues study is an extremely important component of faking literature because it utilized behavioral observation. Although there have been many studies examining risk assessment measures in regard to faking, few studies have specifically examined the association between behavioral observation and faking on self-report measures. But, a limitation for the present study is that the authors used the CTSPC over the CAP as a risk assessment measure.

Although there have been important studies examining faking good on the CAP, the research base remains limited. This differs from the abundance of studies utilizing the MMPI-2 in parental fitness and child custody evaluations. Carr and colleagues (2003) examined socially desirable responding, but in a different context than the researchers in the Budd and colleagues (2000) study. Parent participants (91 mothers and 73 fathers) were recruited from the Family Court Centre, an agency which primarily administered court-ordered parental capacity assessments (e.g., assessments conducted to establish minimal parenting standards acceptable to maintain parental rights of a child) (Carr et al.). Therefore, these parents may have felt more pressure to fake-good because they had open CPS cases. Parents were administered the CAP, the MMPI-2, and the Personality Assessment Inventory (Morey, 1996). All of these measures have validity scales to address socially desirable responding. Additionally, they were administered the
Child Behavior Checklist (CBCL; Achenbach, 1991) to examine differences in parent report of child behavior.

The percentage of invalid profiles in this sample was much greater than general population parents (7-11%; see Milner, 1986) and the adolescent mothers in the Budd and colleagues (2000) study (19%). Forty-nine percent of CAP profiles were invalid, and all were due to an elevated Faking-good index. Examination of the MMPI-2 utilizing a more conservative T-score of 70 (as compared to a score of 65) also revealed 49% of invalid profiles, with response distortion greatest on the L scale (Carr et al.). A lower percentage of parents (17.2% for mothers and 18.2% for fathers) had an elevated score on the Positive Impression Management subscale on the Personality Assessment Inventory. Finally, biological mothers portrayed their children with significantly better behavior (as measured by the CBCL) than foster parents and teachers of the same children (Carr et al.). Notably, the validity indices on the CAP and MMPI-2 were highly correlated, and were comparably able to predict socially desirable responding.

Parental intelligence was controlled in this study by individually assessing each parent and obtaining intelligence quotient scores using Weschler scales (Psychological Corporation 1997; 1999). The high percentage of invalid profiles across mothers and fathers in the study remained similar even when average intelligence ($M = 99.6, SD = 12.1$ for fathers and $M = 92.1, SD = 16.7$ for mothers) for the majority of participants was demonstrated. Therefore, this study did not find the same difference in testing ability as compared to the adolescent mothers studied by Budd and colleagues (2000). One important difference across both studies is the context in which the parents were recruited. Parents in the Carr and colleagues (2003) study were court-ordered to receive a parental fitness evaluation, and already had involvement in a family court for inappropriate parenting practices. The assessment setting could account for the much
higher amount of invalid profiles, given the different levels of motivation experienced by the parent participants. Parents who were already involved in an open CPS case may have felt more motivation to respond in a socially desirable manner.

Another important difference was the test administration between the two studies. Parent participants were administered all measures orally by Budd and colleagues (2000); the administration procedure in the Carr and colleagues (2003) study was not specified. More likely than not, parents were administered the assessments as paper-pencil measures. Given the high amount of parents who fell into either below-average intelligence (i.e., an IQ score of 71-90 for 31 parents) or even mild mental retardation (i.e., an IQ score of 70 or below for six parents) (Carr et al.), administering the self-report measures via pencil-paper assessments could have affected comprehension of each item. Individual reading ability was not measured in this study, which can differ from parental intelligence based on access to education and other resources. The percentage of parents with invalid profiles may have looked more like the mothers in the Budd and colleagues study if reading achievement was compared. Finally, Carr and colleagues failed to include other important measures such as behavioral observation, parental psychopathology, substance abuse, and even a social support measure (included in the study by Budd et al.). The lack of behavioral observation data in both studies is notable considering that parent-child behavioral observations have been cited as important assessments in parental fitness evaluations.

Studies have also been done using other risk assessment measures like the Parenting Stress Index (PSI; Abidin, 1995). The PSI includes a Defensive Responding scale used to detect response distortion. A study done by Acton (1989) assessed deception in child abusers who either had a substantiated report or who were self-defined abusers. In this study, the PSI was able to differentiate between respondents who did and did not engage in deception. Acton (1989)
compared the scores obtained on the PSI to the CAP to accurately differentiate between both groups.

LaFiosca and Loyd (1986) examined socially desirable responding utilizing clinical (n = 39) and non-clinical (n = 22) populations. In this study the target individual was the child. Therefore, no data were presented as to whether or not mothers in the study had a history of maltreatment. Each participant was administered the PSI, the State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970), and the Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1960). The Marlowe-Crowne Social Desirability Scale was used as a measure of defensiveness, and mothers were divided into low-defensive, moderately-defensive, and high-defensive groups. The authors found a positive relationship between high-defensiveness on the Marlowe-Crowne Social Desirability Scale and reporting parenting stress on the PSI. Additionally, mothers with high defensiveness scores were able to mask their responses on the State-Trait Anxiety Inventory, appearing more like mothers with low levels of anxiety. Interestingly, there were no differences between the clinical and non-clinical groups. It was concluded that a high level of defensiveness could potentially distort future mental health recommendations, and distorts the true clinical presentation of behavior problems. As this study did not include measures assessing abuse history or potential, it is hard to draw conclusions to a physically abusive population.

Additional studies have been conducted focusing solely on the MMPI-2. Similar to the Carr and colleagues (2003) study, Ezzo, Pinsoneault, and Evans (2007) examined differences between maltreating parents presenting at risk for termination of rights and parents involved in non-termination child custody hearings. Seventy maltreating parents were included, and a control group of 205 non-maltreating parents involved in child custody hearings were recruited from a
juvenile court setting. All MMPI-2 clinical and validity scales were examined to assess for differences between the two groups. Although the child maltreatment group presented differently on many of the clinical scales (e.g., psychopathic deviate, paranoia, schizophrenia, hypomania, and social introversion), the authors reported that there were no significant differences in the frequency of invalid profiles for the maltreating and non-maltreating groups. In fact, the maltreating group had significantly higher elevations on the faking-bad scale (i.e., F scale) on the MMPI-2 as compared to controls. Although this study was able to detect differences between the two groups, it did not show a difference in defensive responding. Maltreating parents were no better at masking psychopathology, despite being involved in a high-stress, high-risk context.

Similar findings were reported in a study by Posthuma and Harper (1998), in which no differences were found on the L scale of the MMPI-2 (Butcher et al., 2001) between parents involved in custody evaluations for maltreatment (57% of the sample; this included both physical and sexual abuse) and those involved in evaluations for non-maltreatment. The authors concluded that the L scale on the MMPI-2 may not be appropriate for “unsophisticated” individuals (Posthuma & Harper, pp. 439), but did not include any measures of intelligence, years of education, parental substance abuse or psychopathology. Therefore, it is difficult to understand how the authors defined “unsophisticated.” Despite the lack of additional measures, it appears that the potential consequences of child custody evaluations (e.g., termination of parental rights) did not necessarily change responding for parent participants in this study. Both of these studies demonstrate the mixed results from past research studies examining faking behavior, and the need for a multimethod approach to assessment.

The MMPI-2 has also been widely used in studies examining child custody evaluations outside of maltreatment. One notable study conducted by Bathurst, Gottfried, and Gottfried
(1997) examined MMPI-2 profiles of 508 parents involved in court-ordered child custody evaluations. No data were provided as to whether or not parents involved in the evaluations had a history of child abuse or neglect. Parents were also better educated than parents from other faking studies (e.g., average level of education was “some college”; Bathurst et al., 1997), but intelligence was not assessed. Parental psychopathology was assessed by examining the clinical scales on the MMPI-2. Parents in the sample scored much higher on both the L and K scales as compared to the standardized norms on the MMPI-2. Interestingly, when comparing defensive versus non-defensive parents, the authors found no differences in regard to psychopathology between the two groups, concluding that parents involved in child custody evaluations (regardless of defensiveness) did not possess pathological behaviors. It was expected that parents in this sample would feel motivated to respond defensively, but what is interesting is that this motivation remained even though the consequences from being involved in the evaluation were most likely less serious than maltreating parents (e.g., change in parental rights versus termination of parental rights). Also, parents in this study were more educated than samples in similar studies, leading to mixed conclusions on the importance of level of education in faking behavior. The authors’ conclusion regarding parental psychopathology appears to be too generalized, especially given that the MMPI-2 was the only assessment used to measure this variable.

Bagby, Nicholson, Buis, Radovanovic, and Fidler (1999) attempted to replicate the findings from the Bathurst and colleagues (1997) study and explore the issue of parent psychopathology using the L and K scales as well as “nontraditional” validity scales (e.g., the Superlative scale, see Butcher & Han, 1993, and the Wiggins Social Desirability scale, see Wiggins, 1959). In contrast to the Bathurst and colleagues study, using the “nontraditional”
Validity scales yielded greater accuracy in identifying faking versus non-faking parents. Additionally, the authors deviated from the outcomes of the Bathurst and colleagues study by finding that a large proportion (40%) of their sample was still elevated on clinical scales (e.g., Paranoia and Psychopathic Deviancy). The authors were still unable to draw any conclusions about the differences between the faking and non-faking groups, but emphasized the need for a larger and more diverse battery of assessments in future evaluations (Bagby et al.).

An important commentary by Medoff (1999) and study by Cooke (2010) highlighted the methodological issues from previous studies examining the L and K scales on the MMPI-2. The contributions by both authors should be replicated with studies involving the CAP validity scales because they addressed demographic variables (e.g., intelligence and socioeconomic status) and clinical versus statistical validity issues when utilizing the validity scales on the MMPI-2. Medoff specifically addressed clinical versus statistical validity when using the MMPI-2 in child custody evaluations, stating that many of the invalid profiles from past studies (e.g., Bathurst et al., 1997; Bagby et al., 1999) represented statistical significance but not clinical significance. He argued for the need to comprehensively assess an individual’s personality and behavioral characteristics outside of the context of a child custody evaluation. He cautioned individuals who interpret MMPI-2 scores to consider the high-stress context of child custody evaluations as a function of defensiveness, outside of a parent’s “normal” behavior.

Cooke (2010) addressed the impact of parental education level and socioeconomic status using 50 parent participants on the ability to respond defensively on the MMPI-2. He cited the lack of demographic information reported from past research (e.g., Bathurst et al., 1997; Bagby et al., 1999), and concluded that differences in education and SES may alter responding on self-report measures. He also attempted to control for context (e.g., defensive responding as a
function of feeling judged or evaluated by mental health professionals) by administering the MMPI-2 to parents in a quiet, private setting. In his study parents had a high level of education ($M = 17.0$ years of education for fathers, $M = 16.87$ years of education for mothers), similar to the sample obtained by Bathurst and colleagues (1997). Additionally, the sample was inferred as having a high socioeconomic status as evidenced by 28% of fathers and 20% of mothers having post-graduate degrees such as law or medical degrees, or a medical residency. Cooke found that more educated parents were able to fake more than lower educated parents in other samples and argued for a higher validity scale cutoff based on parental education. Most importantly, he argued that parental responses on the MMPI-2 should be compared to observed behaviors, information from clinical interviews, and data from other assessment measures to develop patterns of individual behavior.

The large amount of research studies on defensive responding using the MMPI-2 in child custody cases is impressive, but leads to mixed conclusions. Some studies found that setting (e.g., whether or not the parent was involved in a court-ordered child custody evaluation) may have been a variable which engendered defensive responding. Other studies found no differences between parents involved in child custody evaluations based on outcome (e.g., termination of parental rights for child maltreatment versus modification of parental rights due to divorce). Overall, most of the studies did not include other salient parental characteristics, which may have helped to define what defensive parents look like. Similar to the studies done by Budd and colleagues (2000) and Carr and colleagues (2003), it would have been important for these studies to include demographic measures, intelligence and reading comprehension assessments, measures of parent psychopathology (outside of using the MMPI-2 clinical scales), and a behavioral observation component. Additionally, demand characteristics (e.g., average amount of
time it took to complete the MMPI-2; potential experimenter bias) may have been helpful in attempting to define defensive parents.

Therefore, Medoff’s (1999) commentary and especially Cooke’s (2010) study are extremely important additions to the research base on defensive parents. These studies appear to be some of the first to attempt to describe and differentiate faking parents from non-faking parents. They moved away from simply examining the behavior of faking good on the MMPI-2, to exploring setting variables (e.g., the context of the child custody evaluation), and individual parental characteristics involved in faking. Interestingly, Cooke’s study found that highly educated parents were able to fake on the MMPI-2, differing from less educated parents in other samples. These results differ from studies utilizing the CAP, in which parents with less education lower reading comprehension, or lower intelligence either performed the same or worse than parents with average education, reading comprehension, and intelligence. However, the differences in the sample and in the evaluation setting (e.g., private practice versus child protective services agency) may be too large to draw definite conclusions. Therefore, it is important that similar studies are conducted utilizing the CAP.

Even though there has been extensive research focusing on socially desirable responding using a variety of risk assessment measures (e.g., CAP, PSI) and personality inventories (e.g., MMPI-2), research specifically focusing on the faking-good index on the CAP is limited. Because of its longstanding use in child custody evaluations and parental capacity assessments (Carr et al., 2003), the MMPI-2 appears to have the largest research base. Studies examining the CAP have been strong (see Budd et al., 2000; Carr et al., 2003), but have opened up a multitude of exploratory questions, identifying the need for more research examining the behavior of parents classified as faking-good on the CAP.
Treatment Outcomes of Faking-good Parents

Not only is it important to describe faking parents, but it is also important to consider treatment outcomes of parents who initially present as faking-good. Many research studies have focused on the antecedents (e.g., instructional cues; see Milner & Crouch, 1997) or the topography (e.g., Bennett et al., 2006; Budd et al., 2003; Carr et al., 2000; Ezzo et al., 2007; Posthuma & Harper, 1998) of faking-good behavior. This makes sense given the need to gather information before treatment decisions are made. Given the tendency of past research to focus solely on faking behavior at pre-treatment, there is a lack of literature on differences in treatment outcome for those parents who are still involved in mental health services despite having invalid risk assessment profiles. Some studies are present examining the progression of faking participants through treatment (see Gannon & Polaschek, 2005; Walfish, 2011) but these studies have not focused specifically on parents involved in parental fitness evaluations.

Depending on how parents progress through treatment services, the CAP may have most of its usefulness at pre-treatment, and may be an important marker for expected behavior through therapy services or child welfare involvement. Parents who are more motivated to fake-good during pretreatment assessment may be more motivated to do well in treatment, and therefore may show better outcomes at posttreatment assessments. More research needs to be conducted to assess treatment outcomes of individuals who present with socially desirable responding early in the treatment process.

Statement of the Problem

Parental fitness evaluations for abusive caregivers are high-risk, emotionally-charged situations that are difficult for both the mental health professional and the offending parent (Carr et al., 2005; Otto et al., 2000). The potentially life-altering consequences of parental fitness
evaluations (e.g., significantly reduced visitation, or even termination of parental rights) may engender socially desirable behavior from parents. Although there have been risk assessment measures and personality inventories developed with sound reliability and validity in identifying socially desirable or defensive responding (e.g., the CAP, MMPI-2), mental health professionals still have difficulty in identifying characteristics of parents who fake on these measures. Additionally, many research studies focusing on faking behavior have not included a measure of behavioral observation. Given that behavioral observation has been cited as a necessary component of parental fitness evaluations, the lack of research on the relationship between behavior and faking good on self-report assessments is an important limitation of this literature. Although researchers have begun to examine characteristics of faking-good parents utilizing the CAP (e.g., Budd et al., 2003; Carr et al., 2003; Milner & Crouch, 1997), other risk assessments such as the PSI (Acton, 1989; LaFiosca & Lloyd, 1986), and personality measures such as the MMPI-2 (see Bagby et al., 1999; Bathurst et al., 1997; Cooke, 2010; Medoff, 1999), conflicting results still predominate the literature in differentiating faking-good from non-faking-good parents. Finally, it is important to identify treatment outcomes for faking versus non-faking parents. Validity scales on risk assessment measures may relate differently to pre-treatment and post-treatment concerns. Specifically, at pre-treatment they may be useful in identifying salient demographic variables (e.g., intelligence, level of education, reading ability) and behavioral variables (e.g., motivation to fake) which may become important as treatment progresses. Parents who are motivated to fake at pre-treatment may actually be more highly motivated to do well in treatment, which could affect their skill level and functioning at post-treatment.

This study attempted to address three major aims related to faking-good behavior in physically abusive parents using an exploratory approach. The first aim of the study focused on
exploring the characteristics of the faking-good sample as measured by the CAP using data from a large, national dataset involving physically abusive parents involved in the child welfare system. Additionally, parents in the sample were compared to parents who do not fake-good on the CAP. The second aim was to examine differences in behavioral observation between faking-good and non-faking-good parents. Finally, the study examined differences in treatment outcome to assess if faking-good parents do better or worse when involved in the mental health system.

**Exploratory Questions**

**Exploratory Question 1:**

What are the characteristics of parents with an elevated faking-good index on the CAP, and how do these parents differ from those without an elevated index?

Because the literature has not produced consistent variables that discriminate between faking and non-faking parents, no specific hypotheses were proposed in this area. However, this study incorporated variables used in past research studies such as level of education and IQ. Additionally, parental CPS welfare involvement (e.g., physical abuse referrals, children removed from the home) was examined to address potential differences in faking based on experience with the system. Finally, parents were compared on total scores from a depression measure, antisocial personality disorder symptoms, and substance abuse/dependence to examine differences outside of the context of a parental fitness evaluation.

**Exploratory Question 2:**

What differences, if any, exist between faking and non-faking parents on behavioral observations of parenting?

Given the lack of literature identifying faking-good behavior in abusive parents, and differentiating that behavior from non-faking parents, there were no hypotheses established in
this area. However, differences in positive (e.g., labeled praise, positive physical touch) and negative (e.g., critical statements, negative physical touch) parental behaviors were assessed in this study utilizing the Dyadic Parent-Child Interaction Coding System, second edition (DPICS-II; Eyberg, Bessmer, Newcomb, Edwards, & Robinson, 1994). The examination of both positive and negative parenting behaviors will be used to add to the literature on behavioral observation and faking-good behaviors.

*Exploratory Question 3:*

Do parents who fake-good perform better or worse over time as compared to non-faking parents when administered mental health services, including evidence-based treatments?

Although there are no hypothesized differences between faking-good and non-faking-good parents, future abuse risk (e.g., prospective recidivism rates) and change scores on DPICS parent negative and positive composites were examined.

**Method**

**Participants**

Data utilized in this study were collected from participants recruited through the University of Oklahoma Health Sciences Center by Chaffin and colleagues (2004) as part of a large-scale study examining abuse recidivism and Parent-Child Interaction Therapy (PCIT) (Eyberg & Robinson, 1982; McNeil & Hembree-Kigin, 2010). The study was funded by the United States Department of Health and Human Services, Administration on Children, Youth and Families, Children’s Bureau, Office on Child Abuse and Neglect (Chaffin, 2004). Data collection was conducted from 1997-2001 and participants were recruited from central Oklahoma. The database was made available for secondary data analysis through the National Data Archive on Child Abuse and Neglect through Cornell University.
Parent-child dyads were referred to the study through the child welfare system after a confirmation of a physical abuse report (Chaffin et al., 2004). Three hundred parent-child dyads were originally referred to the study (Chaffin et al., 2004). Inclusion criteria were as follows: (a) the abusive parent was able to participate in the study with at least one abused child without termination of parental rights; (b) the abusive parent had a score of 70 or higher on the Kaufmann Brief Intelligence Test (KBIT); (c) the identified child was between the ages of 4-12 years old; (d) the abusive parent was not involved in the child welfare system due to sexual abuse; and (e) the parent was able to provide informed consent for participation (Chaffin et al., 2004). One hundred twelve parent-child dyads (37% of referrals) met inclusion criteria for the study; during assessment administration two additional parent-child dyads were removed from the study due to parental difficulty with comprehension of assessment measures, leaving a total of 110 participants (Chaffin et al., 2004). Of the total 110 parent participants in the sample, 71 (64.5%) were female and 37 (33.6%) were male; the average parent age was 32.4 ($SD = 8.68$). There were 43 female children (39.1%) and 67 male children (60.9) included the study; the average child age was 8.04 ($SD = 2.76$). All data were collected on the identified abusive parent and abused child; nonabusive parents or parent partners and nonabused children were included in the study as collateral participants, but did not include any study data (Chaffin et al., 2004).

In the Chaffin and colleagues (2004) study, parent-child dyads were administered a baseline assessment. All parent-child dyads received a baseline assessment, which included all variables examined in this study. Parents in each group were assessed using the Dyadic Parent-Child Interaction Coding System, Second Edition (DPICS-II). Behavioral observations using the DPICS-II consist of three 5-minute observations. In the first observation, child-led play, parents are instructed to let their child lead the play, and to refrain from using any commands or
prompts. In the second observation, parent-led play, parents are instructed to keep the child playing according to their rules. In this interaction they are expected to be directive, and to lead the play. In the final observation, child clean-up, parents are instructed to give their child commands to clean up all of the toys in the room without parent help. Each situation is conducted to assess parental skill level in engaging in a positive interaction with their child, in and being able to manage their child’s behavior. Additionally, children are assessed for their level of compliance to parental commands.

Following the baseline assessment, all parent-child dyads were randomly assigned to one of three study groups: Standard PCIT, Enhanced PCIT (EPCIT) which included additional services to PCIT targeting parental depression, substance use, and other family problems, and a standard community program. Forty-two parents (38.2%) were assigned to the PCIT group, while 33 parents were assigned to the EPCIT group (30.0%), and 35 parents (31.8%) were assigned to the standard community program. A total of 26 (23.6%) parents had an elevation on the Faking-good index on the CAP, while 83 (75.5%) parents did not have an elevated score in this sample.

Measures

*Demographic Information.* Demographic data were collected utilizing a measure created for the study. The measure included culturally sensitive items (e.g., for Native American and Hispanic participants) and was available in both English and Spanish (Chaffin et al., 2004). The measure had strong two-week test-retest reliability (α = .74) and interrater reliability (kappa = .79). The measure was organized into four separate sections, collecting information on the background of the participant (e.g., housing and financial status, whether or not parent/family received social service programs, information on race/ethnicity, religious status, marital status,
and level of education), the parent’s lifestyle choices (e.g., seatbelt use, telephone use, drug/alcohol use), social relationships (e.g., church and community involvement), and family relationships. The current study analyzed only the following variables: level of education, and if the parent ever had their children removed from the home.

Child Abuse Risk. Risk for child abuse was measured utilizing the Child Abuse Potential Inventory-IV (CAP; Milner, 1986). The CAP is a 160-item self-report measure used as a screening device for physically abusive and non-abusive parents. It contains six abuse risk factors including: distress, unhappiness, rigidity, problems with child/self, problems from family, and problems from others. It also contains a broadband dimension for identifying child abuse. The measure also contains three validity scales, including lie (containing 18 items), inconsistency and random responding. Combinations of the three validity scales identify additional response distortions including the Faking-good index, the Faking-bad index, and the Random response index. The Faking-good index contains an elevated lie scale and a random responding scale within normal limits, while the Faking-bad index contains an elevated random response scale and an inconsistency scale within normal limits. The Random response index contains both an elevated random response scale and inconsistency scale. Internal consistency on the CAP is strong and ranges from .95 to .86 (Milner, 1986). The CAP also has high test-retest reliability at a one-day interval (.91) to a three-month interval (.75). The present study used only the Faking-good index of the CAP to divide the sample into two groups. The Faking-good sample included parents with either a Lie score of 7 or 8, depending on level of education (see Milner). The non-faking group contained parents with no elevation on the Faking-good subscale on the CAP.
Parental Involvement in the Child Welfare System. Chart reviews were completed on each parent-child dyad to assess past and future abuse reports throughout the duration of the study. All chart reviews were completed through the state of Oklahoma’s computerized child welfare system database. Reports were separated into physical abuse, sexual abuse, and neglect, and research assistants manually checked the data to ensure the report correctly matched the parent participant (Chaffin, 2004). The present study examined both past and future reports of child physical abuse and neglect. Past physical abuse reports were be coded as a continuous variable (e.g., frequency counts of past referrals), and future physical abuse reports were coded as a dichotomous variable (e.g., reoffended versus did not reoffend).

Parent-Child Interactions. To assess the topography and quality of parent-child interactions, the Dyadic Parent-Child Interaction Coding System, Second Edition (DPICS-II; Eyberg et al., 1994) was utilized. The DPICS-II codes verbal, vocal, and physical behavior for both the parent and the child. Verbal, vocal, and physical behaviors are divided into “positive” and “negative” behaviors (e.g., a labeled praise is considered a positive behavior, while a child whine and a parent push or shove are considered to be negative behaviors). The DPICS-II has demonstrated strong interrater and test-retest reliability as well as discriminant validity and treatment sensitivity (see Aragona & Eyberg, 1981; Bessmer, 1998; Foote, 200; Schumann, Foote, Eyberg, Boggs, & Algina, 1998; Webster-Stratton, 1985). In the Chaffin et al. (2004) study, videotaped parent-child interactions were conducted, and then were coded by research assistants unaware of group assignment. Standard criteria were met for each research assistant, including mastery of coding against the trainer before coding any study data, and periodic coder checking by the trainer to assess drift. Additionally, tapes for seven participants were sent to an independent, off-site coder to check inter-rater reliability (Chaffin et al., 2004). Coders had high
interrater reliability, with a correlation of .94 for negative parent behaviors, and .84 for positive parent behaviors (Chaffin et al., 2004). The present study focused solely on pre and post parental verbal and physical behavior composites including a negative behavior composite containing proportion scores of critical statements and negative physical touch in relation to total parent behaviors, and a positive behavior composite using proportion scores of praise and positive physical touch. Appendix A contains examples of critical statements, negative physical touch, praise, and positive physical touch.

Depressive Symptoms. Parental endorsement of depressive symptoms was measured by the Beck Depression Inventory, First Edition (BDI-I; Beck et al., 1961). The BDI is a 21-item self-report measure containing clinical cutoff scores indicating minimal, mild, moderate, or severe depressive symptoms. The BDI has strong internal consistency and test-retest reliability, and strong construct validity. It is a widely used measure in the assessment of depressive symptoms. The present study utilized the BDI pre-treatment total score, as it is an indicator of severity of depressive symptoms.

Substance Abuse. The Diagnostic Interview Schedule (DIS; Robbins, Helzer, Croughan, & Ratcliff, 1981) Alcohol and Drug Modules was modified for use as a parental self-report measure to assess substance abuse in the prediction of potential child physical abuse. The modules were developed based on criteria from the Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DSM-III; American Psychiatric Association, 1980). Both the alcohol and drug modules have strong convergent validity with other alcohol and drug measures (see Chaffin et al., 2004; Goethe & Ahmadi, 1991; Goethe & Fischer, 1995; Hasin & Grant, 1987). The present study utilized codes associated with the presence of alcohol and drug dependence and abuse.
Antisocial Personality Disorder (ASPD) Behaviors. The DIS (Robbins et al., 1981) Antisocial Personality Disorder module was used to assess parental ASPD symptoms. This module was also modified to be used as a parental self-report measure. The present study utilized the total amount of parental ASPD symptoms. ASPD symptoms include: academic difficulties, school expulsions, vagrancy, running away from home, being arrested, lying, stealing, vandalism, fighting, job troubles, being negligent toward children, marital problems, violence, and trouble with driving.

Intelligence. Parental intelligence was measured utilizing the Kaufmann Brief Intelligence Test (KBIT; Kaufman & Kaufman, 1990). The KBIT measures verbal and nonverbal intelligence and is individually administered to participants aged 4 to 90 years. It contains two subscales: Vocabulary and Matrices. The KBIT was included as a screening measure for abusive parents as a marker of inclusion criteria for potential referrals. Additionally it was administered to the abused children (Chaffin et al., 2004). The present study utilized the KBIT total score for the parent participants in analyses.

Analyses

The present study utilized descriptive and experimental analyses to describe faking-good versus non-faking parents. Descriptive analyses included using frequency counts and variable means on multiple parent and child variables. For the experimental analyses, independent-samples t-tests were conducted on all continuous variables. Independent-samples t-tests were considered to be appropriate tests to compare mean differences between the faking and non-faking parents on all continuous variables. Chi-square analyses were conducted on all dichotomous variables. All continuous variables analyzed using independent-samples t-tests were checked for normality by examining skewness and kurtosis. Additionally, homogeneity of
variance was assessed for each variable. A sample size of 109 subjects was deemed to have adequate power to conduct both independent-samples \( t \)-tests and chi-square analyses.

Results

A total of 26 (23.6\%) parents had an elevation on the Faking-good index, while 83 (75.5\%) parents did not have an elevated score. Parents and children in the faking group had an average age of 28.88 \((SD = 7.07)\) and 7.31 \((SD = 2.57)\), while parents and children in the non-faking group had an average age of 32.8 \((SD = 8.18)\) and 8.18 \((SD = 2.76)\). Of the 26 parents with invalid profiles at pretreatment, 8 (32\%) parents were involved in standard PCIT, 8 (32\%) were involved in enhanced PCIT, and 9 (36\%) received the standard community program.

Exploratory Question 1. What are the characteristics of parents with an elevated faking-good index on the CAP, and how do these parents differ from those without an elevated index?

To test the first exploratory question, independent-samples \( t \)-tests were conducted on all continuous variables, and chi-square analyses were conducted on all categorical variables. Faking versus non-faking parents was entered as the grouping variable in analyses. Continuous variables included parent IQ measured by the KBIT, parent depression measured by the BDI, total number of Antisocial Personality Disorder (ASPD) symptoms at pretreatment measured by the DIS, and number of past physical abuse referrals. Categorical variables included parent level of education, and presence/absence of any alcohol and any drug diagnosis. All continuous variables were checked to meet assumptions of an independent-samples \( t \)-test. Assumptions include normality of data, homogeneity of variance, and independence of observations. Parent KBIT scores met all assumptions. Parent BDI scores, ASPD symptoms, and past physical abuse
referrals violated assumptions of normality, and were all square-root transformed. Once transformed, all variables met assumptions, and were used in analyses.

In regard to testing ability, there were significantly different levels of IQ for faking ($M = 91.40, SD = 10.13$) and non-faking parents, ($M = 96.61, SD = 10.03$); $t(106) = 2.27, p < .05$, yielding a small effect size ($\eta^2 = .05$). There were no significant differences between faking and non-faking parents on level of education, $\chi^2 (6, N = 107) = 5.98, p > .05$; parents in the faking group on average completed a high school GED while parents in the non-faking group, on average, completed a high school diploma. Parents did not differ in regard to child welfare involvement. There was not a significant difference between groups on past physical abuse referrals, $t(107) = .98, p > .05$. Parents in both groups had an average of 2 past physical abuse referrals. Additionally, there was not a significant difference between groups on whether or not the child had ever been removed from the home, $\chi^2 (1, N = 102) = 1.09, p > .05$. Eleven (42.3%) parents in the faking group and 25 (30.1%) parents in the non-faking group had their children previously removed.

In regard to psychopathology at pretreatment, there were significantly different levels of depression for faking ($M = 6.79, SD = 7.43$) and non-faking parents ($M = 13.91, SD = 10.41$); $t(104) = 3.62, p < .001$, yielding a moderate effect size ($\eta^2 = .11$). Parents in the faking group ($M = 5.00, SD = 2.40$) did not significantly differ in number of ASPD symptoms at pretreatment from parents in the non-faking group ($M = 4.66, SD = 2.74$); $t(107) = -.75, p > .05$. BDI and ASPD symptom means reported for both groups were obtained before the data were square-root transformed. Additionally, no significant differences were found between faking and non-faking parents on having a diagnosis of alcohol abuse/dependence, $\chi^2 (1, N = 109) = .61, p > .05$, or drug abuse/dependence, $\chi^2 (1, N = 109) = 1.31, p > .05$. A small percentage of parents in both groups had a substance abuse diagnosis at pretreatment: 3 (11.5%) of faking parents and 15 (18.1%) of
non-faking parents had a lifetime diagnosis of alcohol abuse and 3 (11.5%) of faking parents and 18 (21.7%) of non-faking parents had a lifetime diagnosis of drug abuse at pretreatment. Table 1 displays results from all results examining differences in parental characteristics.

**Exploratory Question 2. What differences, if any, exist between faking and non-faking parents on behavioral observations of parenting?**

There were 17 out of 109 cases (15.5%) missing behavioral observation data utilizing the DPICS-II. Therefore, a Missing Values Analysis was performed on the missing items to assess any patterns of missingness. According to Tabachnik and Fidell (2007), variables are considered Missing Completely at Random if they have a nonsignificant Little’s MCAR test. Unfortunately, the data yielded a significant Little’s MCAR test, \( \chi^2 (12, N = 92) = 25.36, p < .05 \), so this was a violation. Variables are considered Missing at Random if they show that missingness is not related to the dependent variable. All variables showed no statistically significant relation to the dependent variable, with the exception of the DPICS positive behavior composite and whether or not the parents ever had their children removed from the home, \( t (16.8) = 3.5, p < .05 \). Given that there was only one violation, and a large amount of data missing from behavioral observations, the 17 cases were dropped from analyses, and the remaining data were analyzed utilizing independent-samples \( t \)-tests.

Negative and positive behavior composites were analyzed as proportion scores to account for the violation of normality in the data. Behavioral observation data is often positively skewed at pretreatment because parents usually engage in a low frequency of behaviors. Negative composite scores (e.g., critical statements and negative physical touch) and positive composite scores (e.g., praise and positive physical touch) were divided by the total number of parent verbal and physical behaviors across all three pretreatment situations (e.g., child-led play, parent-led play, child clean-up). Proportion scores revealed the total number of either negative or positive
parental behaviors in relation to the total number of parent behaviors. Despite converting
negative and positive composites into proportion scores, the data still violated assumptions of
normality, and thus, had to be square-root transformed.

The square-root transformed data were analyzed using independent-samples t-tests. There
was not a significant difference between parents in the faking and non-faking groups for negative
behaviors at pretreatment, between parents in the faking group ($M = 9.47, SD = 8.97$) and non-
faking group, ($M = 7.51, SD = 5.64$); $t(91) = -.48, p > .05$. Additionally, there was not a
significant difference between parents in the faking ($M = 3.47, SD = 3.64$) and non-faking ($M =
4.15, SD = 6.82$) groups for positive behaviors at pretreatment, $t(91) = -.25, p > .05$. Table 1
demonstrates outcomes in parental behaviors. Finally, bivariate correlational analyses were
conducted to examine potential covariates with behavioral observation data. Parent IQ,
depression score, total ASPD symptoms, and number of past physical abuse referrals were all
entered into a bivariate correlational matrix with the negative and positive behavior composite
proportion scores. No significant correlations were obtained from these analyses, yielding a lack
of covariates in this sample.

**Exploratory Question 3. Do parents who fake-good perform better or worse over
time as compared to non-faking parents when administered mental health services,
including evidence-based treatments?**

A chi-square analysis was conducted to assess recidivism rates between faking and non-
faking parents. Nine (25.0%) parents in the faking group recidivated posttreatment, while 16
(22.7%) parents in the non-faking group recidivated. No significant differences were found
between faking and non-faking parents on future abuse recidivism rates, $\chi^2 (1, N = 108) = .10, p
> .05$. 
There was a small amount of posttreatment behavioral observation data; only nine parents in the faking group and thirty-six parents in the non-faking group had DPICS data at posttreatment. Therefore, a visual analysis of means was conducted for DPICS change scores from pre to posttreatment. For the negative behavior composite, parents in the faking group had a mean change score of -2.67 ($SD = 4.64$), while parents in the non-faking group had a mean change score of -2.97 ($SD = 5.76$). For the positive behavior composite, parents in the faking group had a mean change score of 13.22 ($SD = 14.80$), and parents in the non-faking group had a mean change score of 10.08 ($SD = 10.93$). Exploratory independent-samples $t$-tests were conducted to assess differences between the two groups. No significant differences were found between faking and non-faking parents, $t (43) = -.15, p > .05$ on negative behaviors. Additionally, no significant differences were found between faking and non-faking parents, $t (43) = -.71, p > .05$ on positive behaviors. Although there was insufficient power to truly determine differences between the two groups, visual analysis and exploratory analyses indicated small differences for both negative and positive behavior composites between faking and non-faking parents. Table 2 reflects posttreatment outcomes from both groups.

Discussion

Parents who engage in faking-good behavior on the Child Abuse Potential Inventory (CAP) at pretreatment were compared to parents who do not engage in faking-good behavior. Comparisons were made between both groups of parents on a number of variables, including demographic information, cognitive testing ability (e.g., IQ and level of education), involvement in the child welfare system, behavioral observation data, and posttreatment functioning (e.g., recidivism rates, behavioral observation data). Overall, there were mixed findings regarding differences between parents on all of these variables.
Parents differed significantly on IQ but not on level of education. Specifically, parents in the faking group presented with a lower IQ than parents in the non-faking group. It is important to note, that although there was significant difference in IQ, the difference was considered small. Differences in cognitive testing ability in this sample are inconsistent from past studies utilizing the CAP. For example, the Carr and colleagues study (2003) found no differences in faking behavior of parents on the CAP when controlling for IQ. Budd and colleagues (2000) found differences between faking and non-faking adolescent mothers on a reading comprehension measure, which was the only difference found between both groups. Unfortunately the Budd and colleagues study did not test participants on IQ, so there is no comparison point in this sample.

One difference between this investigation and previous studies using the CAP may have been the manner in which the pretreatment assessment measures were administered. Budd and colleagues reported that measures were administered orally to subjects, but the administration procedure was not specified in the Carr and colleagues study. It can be assumed that subjects were most likely given measures using a paper-pencil technique. Additionally, the Chaffin and colleagues (2004) study did not clarify as to whether or not the measures were administered orally or via paper-pencil format. Therefore, it is difficult to know whether or not administration style affects faking outcomes for parents in a child welfare sample, and the administration style confusion clouds the results of the importance of cognitive testing ability with a faking sample. Furthermore, the small effect size obtained in this study between parents on IQ level may demonstrate that this variable does not have a consistent role in the ability to differentiate between parents who fake good on measures and those who do not.

Interestingly, parents in this sample also did not differ on measures of psychopathology (with the exception of depression scores), which is consistent with results from the Budd and
colleagues (2000) study. Parents in the faking sample did not differ in regard to antisocial personality disorder symptoms or a diagnosis of either alcohol or substance abuse/dependence. Given that the antisocial personality disorder symptoms included lying, violence, stealing, and getting in “trouble,” it may be inappropriate to classify parents in the faking group in this sample as any more conniving or manipulative than parents in the non-faking group. Parents may not necessarily have been purposefully “faking good” but rather may have either had more difficulty with comprehension of pretreatment measures, or may even have had a more positive outlook on life.

The difference in outlook is best demonstrated by the significantly different Beck Depression Inventory (BDI) scores between the faking and non-faking groups at pretreatment. Parents in the faking group had significantly lower BDI scores than parents in the non-faking group, which was confirmed by a moderate effect size. It is important to note that parents may have engaged in faking behavior on the BDI at pretreatment, given its high level of face validity. Perhaps more interestingly, parents in the faking group may have possessed a slightly higher optimistic attitude when asked to assess their parenting practices and experiences. To examine a difference in attitudes, a visual analysis of differences in individual items on the CAP identified differences in perception of life quality between the two groups. For example, 100% of parents in the faking group agreed with the item “My life is happy” while only 53.0% of parents in the non-faking group agreed. Additionally, 100% of parents in the faking group agreed with the item “I love all children,” with 80.7% of parents in the non-faking group agreeing with this item. Parents in the faking group may have simply demonstrated a “Pollyanna effect,” or “positive thinking” which is the tendency for individuals to be overly optimistic, and to focus on positive over negative outcomes in stressful events (Goodhart, 1985). Parents in the faking group faced
with the harsh reality of being involved in the child welfare system may have chosen to be more optimistic rather than focusing on the negative outcomes of the situation.

In fact, research has demonstrated the benefits of being unrealistically optimistic. A study conducted by Taylor and Brown (1988) demonstrated what is often thought as counterintuitive: individuals with higher levels of depression and lower levels of self-esteem may actually have a more realistic outlook on life, whereas holding self-affirming illusions (i.e., a belief that is more general, pervasive, and enduring than a thought error or thought bias; see Taylor & Brown) actually leads to higher levels of optimism, and thus better self-esteem and mental health outcomes. Additionally, depressed individuals may have lower self-perceptions, or self-perceptions which are more congruent with how others perceive them (see Lewinsohn, Mischel, Chaplin, & Baron, 1980). Given the significant differences in the parents’ BDI scores, it may be appropriate to consider parents in the non-faking group as perhaps experiencing slightly more depressive symptoms, thus having a more realistic outlook on life, and having a self-perception more consistent with how their assessors viewed them.

Perhaps one of the most interesting findings was the nonsignificant difference in both negative and positive parent behaviors between both groups at pretreatment. Although they presented themselves more favorably on the CAP, and may have been more optimistic in their attitudes, parents in the faking group did not demonstrate lower levels of criticism or negative physical touch. In fact, parents in both groups demonstrated very low levels of negative and positive behaviors at pretreatment. Parents in the faking group also did not present themselves more favorably than non-faking parents in regard to positive behaviors. They did not display more praise and/or positive physical touch during behavioral observations. The discrepancy between self-report measures and behavioral observation is not a new phenomenon; in fact,
research has thoroughly demonstrated that individuals often demonstrate poor insight with respect to identifying the rationale behind their responses on self-report measures, and behavioral observation data do not necessarily match responses on self-report measures (see Bem, 1974, Mischel, 1968). Therefore, the discrepancy between both modalities of assessment in this study has important implications for parental fitness evaluations.

“Best practice” guidelines have identified that a multimethod assessment is important when conducting parental fitness evaluations (Budd, 2001; 2005), and results from this study add further evidence in favor of this view. It remains extremely important to include an objective measure of parenting style and parenting effectiveness given the bias that can come from self-report measures. Parents may look more competent at pretreatment if only assessed by self-report measures, but may be just as in need of services as parents who present themselves more realistically. Therefore, a behavioral observation component is especially important to include during evaluations. Behavioral observation may yield a less biased and more realistic view of parenting when compared to self-report measures.

No differences were found between faking and non-faking parents on treatment outcome data as well, and a designation of “faking” at pretreatment may not have been extremely applicable to treatment outcome for parents in this sample. Parents in the faking group did not display lower recidivism rates than parents in the non-faking group. In fact, 9 (25.0%) parents in the faking group recidivated, while 16 (22.7%) parents in the non-faking group recidivated, which show similar percentages across both groups. Therefore, an invalid profile at pretreatment was not necessarily indicative of a difference in recidivism rates at posttreatment. Although there was a small amount of behavioral observation data at posttreatment, a visual analysis of means indicated no differences between parents on the reduction of negative behaviors. Parents in the
faking group did demonstrate slightly higher positive behaviors, $M = 13.22$ compared to parents in the non-faking group, $M = 10.08$, but these differences were small. Follow-up exploratory analyses utilizing independent-samples $t$-tests also indicated no significant differences between faking and non-faking parents on negative and positive behaviors.

**Limitations and Future Directions**

This study has multiple limitations, and warrants a discussion of each. The first limitation is inherent in secondary data analysis. Given that the data used in this study were not directly collected, there was no control over what measures were used and how assessments were administered. Additionally, the amount of missing data on the variables used in this study was not able to be controlled, and led to difficulties in data analysis. The largest amount of missing data occurred in the behavioral observations, which was an integral part of this study. This may have accounted for the lack of differences in negative and positive behavior between faking and non-faking parents, and threatens internal validity.

An additional limitation is the use of the BDI-I (Beck, 1961) in this study. Given that this measure is an older version of the measure currently used (BDI-II; Beck, Steer, & Brown, 1996), changes from the first to second edition may render the results from this study as not being generalizable to current research. Future research should investigate changes in the BDI-II to assess if these hold with faking and non-faking parents. Finally, many of the variables used in this study were positively skewed, which also led to difficulty in data analysis. For example, as mentioned earlier, behavioral observation data using the DPICS were most likely skewed at pretreatment because parents often engage in low frequencies of behavior, especially positive behaviors. Parents involved in the child welfare system may not necessarily know to engage in praise or positive physical touch while being observed, and may feel pressure to use a low level
of critical statements and negative physical touch. Additionally, scores on the BDI and the ASPD module may also have been positively skewed because they were used in a sample that was not being directly studied for depression and antisocial behaviors. Although all of these variables were able to be utilized in this study, it made interpretation of the results slightly more difficult.

Results from this study may have implications for future research and practice. The mixed results from this study and from the Budd and colleagues (2000) and Carr and colleagues (2003) studies underscore the need for more research utilizing the CAP’s Faking Good Index. Unfortunately, it appears to still be difficult to consistently distinguish parents who fake good on the CAP from those with valid profiles across demographic, psychopathology, behavioral observation, and treatment outcome variables. Also, given that this study was one of the first to assess treatment outcomes of parents who respond in a socially desirable manner, more studies are needed to assess the effects of treatment on parents who present as faking at pretreatment. It would be important to assess whether or not validity of the CAP at pretreatment really matters for posttreatment outcomes. If future studies corroborate the outcomes from this investigation and from findings from Carr and colleagues and Budd and colleagues, then it may be demonstrated that it still may be helpful to use the CAP as a measure of parenting attitudes and child abuse potential, without simply identifying the entire measure to be invalid and interpretable due to a positive response bias.

It is also very important to consider the use of behavioral observation during parental fitness evaluations. Outcomes from this study showed that even though some parents were able to present themselves in a socially desirable manner on the CAP, they were not able to maintain these behaviors when observed by a mental health researcher. Given the strong argument that self-report data and behavioral observation data often show discrepancies, it is important for
future parental fitness evaluations to follow the “best practice” multimethod approach as outlined by Budd (2001; 2005). Because parental fitness evaluations may have serious and long-lasting consequences for the parent, child, and other family members, a comprehensive, detailed evaluation appears to be the most effective means of assessment in this situation.

Therefore, results from this study demonstrate that the label of “faking good” at pretreatment may be an important indicator of testing ability or of parent functioning at pretreatment, and should be considered when evaluating parents for physical abuse or neglect, but other considerations should also be made. Results of this evaluation did not support the view that parents who fake good are more manipulative or psychopathic than others, and there was no evidence to suggest that this CAP profile is associated with poorer prognosis with respect to child abuse potential. Both types of parents (e.g., faking versus non-faking) should still be considered for further evaluation, monitoring, and intervention, and should learn how to decrease negative parenting behaviors and how to increase incompatible, appropriate alternatives. It is strongly recommended that future replications utilizing the CAP’s Faking-good index are conducted to corroborate these results. Having a strong research base in the CAP’s Faking-good index can add to information on socially desirable responding used by mental health professionals, and may lead to future child welfare system change over time.
References


Lutz, FL.


Table 1

*A Comparison of Faking-good versus Non-faking parents Across Demographic Information, Psychopathology, Child Welfare System Involvement, and Behavioral Observation at Pretreatment*

<table>
<thead>
<tr>
<th>Assessment Measures</th>
<th>Faking (n = 26)</th>
<th>Non-faking (n = 83)</th>
<th>t</th>
<th>$\chi^2$</th>
<th>p</th>
<th>Partial eta squared</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M (SD)$ or %</td>
<td>$M (SD)$ or %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>High school GED</td>
<td>High school diploma</td>
<td>5.98</td>
<td>.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>91.40 (10.13)</td>
<td>96.61 (10.02)</td>
<td>2.27</td>
<td>.03</td>
<td>.05*</td>
<td></td>
</tr>
<tr>
<td>CPS involvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical abuse refs</td>
<td>2.12 (2.34)</td>
<td>2.04 (2.18)</td>
<td>.98</td>
<td>.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child taken from home</td>
<td>42.3%</td>
<td>30.1%</td>
<td>1.09</td>
<td>.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychopathology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>6.79 (7.43)</td>
<td>13.91 (10.40)</td>
<td>3.62</td>
<td>.00</td>
<td>.11**</td>
<td></td>
</tr>
<tr>
<td>Antisocial symptoms</td>
<td>5.00 (2.40)</td>
<td>4.66 (2.74)</td>
<td>-.75</td>
<td>.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol abuse</td>
<td>11.5%</td>
<td>18.1%</td>
<td>.61</td>
<td>.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug abuse</td>
<td>11.5%</td>
<td>21.7%</td>
<td>1.31</td>
<td>.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavioral observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative behaviors</td>
<td>9.47 (8.97)</td>
<td>7.51 (5.65)</td>
<td>-.48</td>
<td>.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive behaviors</td>
<td>3.47 (3.64)</td>
<td>4.15 (6.83)</td>
<td>-.25</td>
<td>.81</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Variables marked with an “*” indicate a small effect size and “**” indicate a moderate effect size. Additionally, “negative behaviors” include critical statements and negative physical touch. “Positive behaviors” include praise (labeled and unlabeled) and positive physical touch.
### Table 2

*Comparing Recidivism Rates and Behavioral Observation Data across Faking and Non-faking Parents at Posttreatment*

<table>
<thead>
<tr>
<th>Outcome Measures</th>
<th>Faking</th>
<th>Non-faking</th>
<th>t</th>
<th>χ²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recidivism rates</td>
<td>25.0%</td>
<td>22.7%</td>
<td></td>
<td>.10</td>
<td>.75</td>
</tr>
<tr>
<td>Negative behaviors</td>
<td>-2.67 (4.64)</td>
<td>-2.97 (5.76)</td>
<td>-.15</td>
<td></td>
<td>.88</td>
</tr>
<tr>
<td>Positive behaviors</td>
<td>13.22 (14.80)</td>
<td>10.08 (10.93)</td>
<td>-.71</td>
<td></td>
<td>.48</td>
</tr>
</tbody>
</table>

*Note. Negative and positive behaviors represent a change score from pre- to post-treatment.*
Appendix A

Examples of Positive and Negative Parental Behaviors using the Dyadic Parent-Child Coding System, Second Edition (DPICS-II; Eyberg et al., 1994)

<table>
<thead>
<tr>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Behaviors</td>
<td></td>
</tr>
<tr>
<td>Labeled Praise</td>
<td>Praise child for the exact behavior he/she is doing</td>
</tr>
<tr>
<td></td>
<td>Parent says: “I like it when you sit so quietlly at the table!”</td>
</tr>
<tr>
<td>Positive Physical Touch</td>
<td>Engage in warm, nurturing physical touch</td>
</tr>
<tr>
<td></td>
<td>Parent rubs the child’s back or gives the child a hug during play.</td>
</tr>
<tr>
<td>Negative Behaviors</td>
<td></td>
</tr>
<tr>
<td>Criticisms</td>
<td>Negatively commenting on child’s behavior and/or making sarcastic comments</td>
</tr>
<tr>
<td></td>
<td>Parent says: “I don’t like it when you scribble all over the paper – you are ruining the coloring book.”</td>
</tr>
<tr>
<td>Negative Physical Touch</td>
<td>Engage in harsh physical touch with the intention of controlling the child’s behavior</td>
</tr>
<tr>
<td></td>
<td>Parent grabs child by the shirt and pulls him to the table.</td>
</tr>
</tbody>
</table>
Author Note

The data used in this publication were made available by the National Data Archive on
Child Abuse and Neglect, Cornell University, Ithaca, NY, and have been used with permission.
Data from *Physical Abuse Treatment Outcome Project: Application of Parent-Child Interaction
Therapy (PCIT) to Physically Abusive Parents* were originally collected by Mark Chaffin.
Funding for the project was provided by the U.S. Department of Health and Human Services,
The Administration on Children, Youth and Families, Children’s Bureau, Office on Child Abuse
and Neglect (Award Number: 90CA1633). The collector of the original data, the funder,
NDACAN, Cornell University and their agents or employees bear no responsibility for the
analyses or interpretations presented here.