1985

The Use of Parent Data in Measuring the Outcome of Child Psychotherapy

Barry Richard Lindstrom
Loyola University Chicago

THE USE OF PARENT DATA IN MEASURING THE OUTCOME OF CHILD PSYCHOTHERAPY

by

Barry Richard Lindstrom

A Thesis Submitted to the Faculty of the Graduate School of Loyola University of Chicago in Partial Fulfillment of the Requirements for the Degree of

MASTER OF ARTS

April

1985

ACKNOWLEDGEMENTS

I would like to express my appreciation for the time and effort of those who have contributed, directly and indirectly, to the completion of this thesis. I would like to thank Joseph A. Durlak, Ph.D., my director, for his support, encouragement, availabilty and, especially, for his help in editing the many drafts of this manuscript. I would like to thank Dan P. McAdams, Ph.D., for being on my committee and for his support, encouragement, and perspective throughout the process. I would like to thank Frank Slaymaker, Ph.D., for his willing consultation on statistical questions. I would also like to thank my professors, past and present, for their contibutions to my academic development, and my peers for their support and challenge in and out of the academic environment.

I would also like to thank my family for their support and encouragment of all my endeavors. Specifically, I would like to thank my wife, Bonnie, for her constant support and understanding.

VITA

The author, Barry Richard Lindstrom, is the son of Mr. and Mrs. Wallace H. Lindstrom. He was born June 2, 1959, in Denver, Colorado. His elementary education was obtained in the public schools of Denver, Colorado, and San Diego, California. His secondary education was completed in 1977 at John Marshall High School, Milwaukee, Wisconsin.

In September, 1977, Mr. Lindstrom entered Carthage College, Kenosha, Wisconsin. He was graduated with the degree of Bachelor of Arts in Psychology in May, 1981.

In August, 1982, Mr. Lindstrom entered Loyola University, Chicago, Illinois.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

There is a general consensus in the literature on the need for multiple measures of outcome in research on child psychotherapy and related treatment approaches. However, different sources or measures may provide different conclusions about the outcome of treatment. Thus, one problem in using multiple sources or measures of outcome is the intrepretation of discrepant findings. Parents are one important source of information regarding the outcome of child psychotherapy. Mothers' and fathers' reports may be different, however, raising the same issues regarding interpretation that exist when using multiple sources. One of the central issues related to the use of parent data is whether to obtain information from one or both parents and whether to use parent's reports separately or combine them in some manner such as an average parent score. Agreement between parents' reports, or interparent agreement, is thus central to the determination of how to utilize parent data, both in research and clinical settings. Although interparent agreement on diagnostic measures has been investigated, less is known about interparent agreement on measures assessing the outcome of therapy.

Traditionally, there has been an emphasis on maternal over paternal report in research as well as clinic settings. Specifically, three problems exist in the current use of parent data. Much of the outcome research using parent data has relied on data collected only from mothers (e.g., Clement, Fazzone, & Goldstein, 1970). Secondly, many studies have used "parent" data without distinguishing whether this includes data from mothers, fathers, or both (e.g., Dubey, O'Leary, & Kaufman, 1982). Finally, those studies explicitly using both parents have failed to address the issue of agreement within the parental dyad (e.g., Firestone, Kelly, & Fike, 1980). The present review of the literature did not find any outcome studies where interparent agreement was measured. Rather than generalize from studies assessing interparent agreement on diagnostic measures, interparent agreement on outcome measures should be assessed. Before determining whether parents' reports should be used separately or averaged in some manner, a better understanding of the factors affecting interparent agreement on diagnostic and outcome measures is necessary.

The purpose of the present study is to examine methodological issues in the use of parent report data to assess the outcome of child psychotherapy. More specifically this study intends to examine: (1) the effects of using maternal and paternal reports separately and in combination with each other, (2) the factors affecting mothers' and fathers' perceptions of their children's behavior as measured on a behavior checklist, both before and after treatment, (3) the level of agreement

between parents regarding their clinic referred child, and (4) the possible factors related to the child and his or her disorder, the child's parents, and the type of treatment approach employed that might affect interparent agreement before and after treatment. Finally, this study will provide further information about the clinical application of the primary measure utilized, the Washington Symptom Checklist (Wimberger and Gregory, 1968).

CHAPTER II

LITERATURE REVIEW

The Need for Multiple Outcome Measures

There is a dearth of methodologically sound research on the process and outcome of child psychotherapy and related treatment approaches. Moreover, there is comparatively less research on child psychotherapy than on approaches to adult treatment (Barrett, Hampe, & Miller, 1978; Tramontana & Sherrets, 1983). Tuma and Sobotka (1983) suggest the presence of an "increased interest and progress in conceptualizing child psychotherapy research" (p. 418). It remains, however, for this "interest and progress" to be put into action. Tuma and Sobotka (1983) have summarized the current methodological problems in child psychotherapy research into four areas needing attention: "(1) the child and his or her disorder, (2) the therapist and his or her personality, (3) intervention techniques, and (4) outcome measures" (p. 422). The present study is intended to focus on issues relate to the fourth area, outcome measures, specifically regarding the use of parent report data.

The use of multiple outcome measures has been advocated in research on child psychotherapy (Barrett et al., 1978; Cass & Thomas, 1979; O'Leary & Turkewitz, 1978; Tramontana & Sherrets, 1983; Tuma & Sobotka, 1983). A general consensus has emerged that "changes should be

evaluated in as many dimensions obtained from as many sources as possible" (American Psychiatric Association (APA), 1982). Many of the above reviewers have called for the adoption or adaptation of Strupp and Hadley's (1977) "tripartite" model. Strupp and Hadley (1977) suggest the need for outcome measures from society, which would include parents and teachers; the client, in this case the child; and professionals, therapists and counselors. Most research has included therapist ratings as the most frequent outcome measures (Vandenbos & Pino, 1980), with fewer studies measuring parent (e.g., Kissel, 1974; O'Leary, Turkewitz & Taffel, 1973) or teacher perceptions (e.g., Kendall & Finch, 1978; Taylor & Hoedt, 1974) and even fewer assessing the child's perception (e.g., Marvit, Lind, & McLaughlin, 1974).

The need for more than one outcome measure is based on the multidimensionality of change and the inadequacy of any given measure or set of measures alone. Multidimensional perspectives require more than one type of outcome measure from more than one source. "No one class of assessment instruments can adequately reflect change in complex human processes. All types are fallible and worthy in different respects and each may best be employed to complement, clarify and correct for the others" (Johnson & Eyberg, 1975, p. 918). Abramovitz (1976) describes the current lack of use of such complementary outcome measures and the resulting implications of such procedures:

> Outcome batteries have often contained only one or two measures, sometimes completed by the same informant, and the reliabilities and validities of these instruments have typically gone unreported. The fewer the measures and informants used in a study, the more likely its results will be contaminated by the former's inadequacies and the latter's biases. (p. 325)

Although Abramovitz (1976) is specifically reviewing research on group psychotherapy, similar criticisms have been raised in other areas of child treatment research. In order to improve the quality of outcome research, "future assessments of therapeutic outcome must be more comprehensively based and provide a better reflection of the multidimensionality of development and psychopathology in childhood and adolescence" (Tramontana & Sherrets, 1983, p. 443).

The use of multiple outcome measures from various sources mitigates the problem of unidimensional measures discussed above. As described by Strupp and Hadley (1977), "a truly adequate, comprehensive picture of an individual's health is possible only if the three facets of functioning- behavior, affect and inferred psychological structure- are evaluted and integrated" (p. 196). These three facets correspond to outcome measures obtained from the three sources of the tripartite model: society, clients, and the therapist, respectively. Although the use of multiple outcome measures more clearly assesses various dimensions of possible therapeutic improvement, choosing and interpreting the results of multiple measures is difficult. The issues related to choosing multiple measures will be discussed first.

Because there are advantages and disadvantages to each type of outcome measure (Atkeson & Forehand, 1978; Johnson & Eyberg, 1975; O'Leary & Turkewitz, 1978) the choice of measures must be based on additional considerations. O'leary and Turkewitz (1978) suggest that "the choice of dependent measures and the data sources will be dictated by

one's theoretical orientation and the particular question of interest" (p. 752). Similarly, according to Abramovitz (1976), "outcome criteria should be relevant to the presenting problems and the goals of the treatment approach under scrutiny" (p. 325). However, an overreliance on outcome criteria or measures unique to each study would make comparisons between studies even more difficult. Therefore, an additional criteria for choosing measures should be relevance to the research literature.

The use of multiple outcome measures presents difficulties in interpretation. One such difficulty is the determination of how much weight to assign each source or measure. This is especially difficult when multiple measures lead to discrepant findings. "The same individual may be simultaneously judged as mentally healthy or mentally ill and, correspondingly, his therapeutic experience may be judged as positive or negative depending on who is evaluating the patient" (Strupp & Hadley, 1977, p. 196). Thus, one of the major difficulties in using multiple outcome measures is how to interpret different sources of outcome data that provide differing results about the outcome of therapy. Differences between multiple measures or raters may be due to factors related to the source, content, or measurement methodology of the outcome measure (Mintz, Luborsky, & Christoph, 1979).

Currently, there is debate over the amount of agreement to be expected among multiple outcome measures or multiple observers ratings. It has been argued that it is probably unreasonable to expect a high degree of agreement:

> Because observers and situations inevitably affect childrens' behavior it is probably more profitable to determine which observers ratings are most predictive of other important characteristics, than to strive for high agreement among diverse observers. (Achenbach & Edelbrock, 1978, p. 1296)

There is disagreement, however. Mintz et al. (1979) suggested that "contrary to a common opinion consensus measures of psychotherapy outcome could be meaningfully defined" (p. 319) and reported "substantial agreement among viewpoints about broadly defined treatment outcomes, although distict viewpoints clearly exist" (p. 319). The Mintz et al. (1979) study was on the outcome of adult psychotherapy but nevertheless demonstrated the effects of different methods of defining and measuring outcome on the level of agreement among sources. Aside from the probable success of obtaining agreement among multiple sources is the question of the usefulness of such attempts. "Obtaining agreement among diverse observers is perhaps less important than determining which sources of observation reveal stabilities that are in turn related to etiology, prognosis and effectiveness of possible treatment appproaches" (Achenbach & Edelbrock, 1978, p.1289). In summary, the use of multiple sources and measures of outcome provides a broader base of assessment but complicates the interpretation of results.

The Use of Parent Data

Parents are an important source of information regarding their children's behavior and represent one source from the tripartite model. As "significant others" (Davidson & Davidson, 1983; Fiske, 1975) parents are an important and advantageous source of information regarding the

outcome of treatment. Parents can provide unique and particularly relevant information unavailable from other sources. In child psychotherapy especially, the use of reports from parents is important because the child is generally referred by a "significant other" for evaluation or treatment. The use of data from significant others is not without its disadvantages, however.

> Chief among the drawbacks are the availability of a knowledgeable cooperative relative or friend, the motivational or perceptual biases of the informant that may color his or her report...and and the interpretive problems of evaluating outcome when that appraisal of the significant other does not converge with the judgements of other interested parties - that is, the patient and the therapist. (Davidson & Davidson, 1983, p. 595)

Although the information provided by parents may be unique and valuable it may also be particulary susceptible to bias. Issues related to the validity of parents' reports will be discussed in greater detail below.

Parent data have been used frequently at the diagnostic end of the child treatment continuum, in both clinical and research settings. Achenbach and Edelbrock (1978) stressed the importance of using parent data, and suggested that parents' "reports should be systematically integrated into classification procedures whenever possible" (p. 1290). How to best integrate these reports has not yet been determined. Moreover, parents' reports are of no less importance or value in assessing treatment outcome. Such applications have been less frequent than the utilization of parent data for diagnostic purposes, however, and the specific issues related to the use of parent data as an outcome measure have remained largely unexamined.

More importance has typically been placed on obtaining information from mothers than from fathers. This practice is probably attributable to several factors involving theoretical considerations and convenience rather than empirical findings. One possible cause of this onesidedness is the predominantly psychodynamic orientation of most early child guidance centers, which emphasized the child's early development in relation to his or her mother. Theoretical considerations do not appear to be the only determinant of this over reliance on maternal reports , however, because Psychodynamic theories also stress the father's importance in the child's development in relation to Oedipal conflicts. Pragmatic considerations have also influenced clinic and research data collection practices. The cost and difficulty of obtaining information from both parents has often been prohibitive; if a child is only accompanied to the clinic by one parent it is usually the mother. Researchers have also reported difficulty in obtaining information from fathers (e.g., Irwin, Levy, & Shapiro, 1972).

In response to the overreliance on maternal report, many researchers have stressed the need for additional data from fathers. Novick, Rosenfeld, Bloch, & Dawson (1966) stressed the need to obtain information from fathers, even though they suggested that mothers may be the best <u>single</u> source of information. Of all the valid items (as judged independently) reported by a least one source, mothers reported 63%, fathers 55%, teachers 22%, home observers 14%, and school observers 12%, (Novick, Rosenfeld, Bloch, & Dawson, 1966, p. 233). Thus fathers are an

important source of additional information and can supplement information obtained from mothers. The current increase in the use of various parent education and training programs as a treatment approach has placed additional emphasis on the role of fathers in the treatment process (Firestone, Kelly, & Fike, 1980; Horton, 1984)

The emphasis on maternal over paternal report has been true in research as well as in clinic settings. Specifically, three problems exist in the current use of parent data. Much of the outcome research using parent data has relied on data collected only from mothers (e.g., Clement, Fazzone, & Goldstein, 1970; Forehand, Griest & Wells, 1979; Kissel, 1974; Lessing, Black, Barbera, & Seibert, 1976; Novick, 1965). Secondly, many studies have used "parent" data without distinguishing whether this includes data from mothers, fathers, or both (e.g., Dubey, O'Leary, & Kaufman, 1983; Fine, Knight-Webb, & Breau, 1976; Leventhal & Weinberger,1975; Lundeen, 1977; Szapocznik, Kurtines, Foote, Perez-Vidal, & Hervis, 1983; Zold & Speer, 1971). Finally, those studies explicitly using both parents have failed to address the issue of agreement within the parental dyad (e.g., Firestone, Kelly, & Fike, 1980; Kogan & Gordon, 1975; Wimberger & Millar, 1968). The present review of the literature did not find any outcome studies where interparent agreement was measured. Rather than generalize from studies assessing interparent agreement on diagnostic measures, interparent agreement on outcome measures should be assessed. This issue will be discussed in greater detail below.

Methods of Obtaining Parent Data

Parents have provided information about their children in a variety of ways. Traditionally such information has been obtained from interviews. Another major form of data collection from parents is the use of a variety of behavioral checklists, rating scales and questionnaires (see Goldman, L'Engle-Stein, & Guerry, 1982; and Humphreys & Ciminero, 1979, for a review of the most frequently used measures of this nature). Behavior checklist data have been one of the most common sources of parent report outcome measures. Outcome studies have included a diversity of rating scales or checklists; with change in scores over time used as a measure of treatment effects (e.g., Firestone, Kelly, & Fike, 1980; Jesness, 1975; Taylor & Hoedt, 1974). Another source of parent outcome data has been the use of general questions about their feelings and attitudes towards the treatment received and its effectiveness (e.g., Kissel, 1974; Leventhal & Weinberger 1975). Parents have also been used as observers and recorders of their childrens' behavior, usually in more behaviorally oriented treatments (e.g., Colletti & Harris, 1977; Eyberg & Johnson, 1974; Patterson, Cobb, & Ray, 1973). The present study focuses on parent data obtained from behavior checklists, specifically the Washington Symptom Checklist (WSCL; Wimberger & Gregory, 1968).

Parent report measures, in the form of rating scales, checklists, and questionnaires have both strengths and weaknesses. The advantages associated with this type of measure are that they are less costly and

time consuming, and may yield more objective and reliable data than other means such as interviews or projective tests (Edelbrock, 1983). The disadvantages of such parent report measures are related, primarily, to the characteristics of the informant. Most specifically, with parents as informants there are the potential problems of reactance, social desirability, demand characteristics and response sets on obtained ratings. When relying on parent report it is difficult to determine whether parents' perceptions, the child's behavior, or both are actually being measured. "Behavioral ratings reflect not only the characteristics of the child, but also diverse and complex characteristics of the informant." (Edelbrock, 1983, p.298). Moreover, many of these measures have shortcomings in available normative data, adequately demonstrated reliability and empirical validation (Humphreys & Ciminero, 1979).

Determining the validity of parent reports on behavior checklists has been approached in terms of construct and criterion validity. Support for the construct validity of parent reports on behavior checklists has involved assessing the discrimination of clinic from non-clinic children on the basis of checklist scores (e.g., Ferguson, Partyka, & Lester, 1974; Sines, Paulker, Sines, & Owen, 1969; Speer, 1971; Wimberger & Gregory, 1968). Humphreys & Ciminero (1979) suggested that this method of establishing validity is confounded, however, because parents are often the referral source and thus the determinant of clinic status. Additional support for the construct validity of such checklists comes

from studies showing their ability to discriminate among diagnostic groups (e.g., Kazdin and Heidish, 1984), and the convergence of checklist measures and clinically derived records, diagnoses, and other observational measures (e.g., Kazdin and Heidish 1984; Thompson & Curry, 1983).

In terms of criterion validity, the question of whether parent reported change relates to actual behavioral change on the part of the child or only to changed parental perceptions is unresolved. The issue is complicated by "the lack of independent criteria for categorizing children (which) makes it difficult to establish criterion referenced validity" (Achenbach & Edelbrock, 1978, p. 1275). Criterion validity can be established by using either the parents' perceptions or the child's behavior as the criterion. Currently, there is disagreement about which criterion to use, and hence, about the validity of parent report measures. Support for the validity of such measures has been argued on the basis of the importance of parental perceptions in and of themselves (e.g., Kissel, 1974; Zold & Speer, 1971). Arguments against the validty of parent report data have been based on the low correspondence between parent reports and direct observation of the child's behavior or reports from other sources. The central issue here is not the low correspondence itself, but the implications of such disagreement between measures and sources of outcome data. These implications are that: (a) parent report should not be used as the sole measure when change in the child's behavior is the goal of treatment, (b) parent

reports can be used as one of a set of measures to assess the cognitive aspects of the problem, and (c) change in parents' perceptions is a valid treatment goal (Humphreys & Ciminero, 1979). Ross (1978) summarized the issues related to the validity of parent report data:

> While parental judgement of the severity of a problem usually serves as the basis for initiating treatment, that judgement at or after termination of treatment may be influenced by factors other than child behavior itself...Yet, since the parents' view of the child's behavior will, in itself, influence that behavior it is certainly important to assess their opinion. On the other hand...evaluation of treatment outcome must include more than parental judgement. (p.614)

In conclusion, the criteria used to establish the validity of parent report measures may vary and the importance of parental versus external ratings of child's behavior change appears to be relative to a given study, depending on the focus of treatment and intended outcome effects. Ideally, conclusions about the effects of treatment should not rely on any one source, be it parents or therapists, and measures should be obtained both from parents and independent sources. Moreover, variability within each source of data needs to be examined; interparent agreement, especially, needs to be discussed.

Analysis of Parent Data

Given the importance and value of obtaining information from parents about their children's behavior, the first issue that needs to be resolved is how to treat parent data. Should we obtain information from both parents, and if we do, how should we use this information? Are mother's and father's reports one piece of information or two sepa-

rate but related pieces? Stated differently, the question is: "Do mother's and father's reports agree with each other enough to be considered jointly or are they discrepant enough to be considered separately?"

Currently, there is disagreement about how to treat parent data. Some researchers have advocated analyzing mother's and father's reports separately, due to the expectation of differences between parents' reports; others have suggested that such differences are not to be expected, thus eliminating the need for separate analyses. Data supporting the separate analysis of parents' reports comes from several studies. Based on moderate inter-parent correlations (.32 to .74) on checklists Guerney, Shapiro, and Stover (1968) concluded that "reports of mothers and fathers should not be regarded as equivalent or interchangeable" (p. 222). Similarly, Ferguson et al. (1974) conluded that the responses of "mothers and fathers should be considered separately" (p. 179) based on different factor loadings for mothers and fathers on various behavioral dimensions. Thus, while mothers' and fathers' reports may be moderately correlated they may differ in important ways based on each parents' unique perceptions of the child. Moreover, Jacob, Grounds, and Haley, (1982) provide a caveat:

> The unsystematic use of data from either parent will significantly increase error variance, which in turn will yield less precise and interpretable classifcations of experimental samples. Of equal importance, the use of only one parent's data could mask important relationships that exist between the reports of one parent (but not the other) and processes and outcomes relevant to childhood psychopathology....It would certainly seem prudent for researchers to obtain data from both parents and to use each parent's ratings of severity and nature of child disturbance in separate analyses. (p.607)

On the other hand, Thompson and McAdoo (1973) suggested that "significant differences between mother's and father's average ratings of clinic boys and girls are not to be expected" (p. 388). This conclusion implies that parents reports might be combined. It must be noted, however,that Thompson and McAdoo's (1973) conclusions are based on average ratings. The use of measures averaged over all group members might obscure important information about variability within individual parent dyads. Therfore, even if significant differences between mothers' and fathers' average ratings may not be expected, it should not be concluded that their reports are similar enough to be combined in individual cases.

An additional argument for obtaining information from both parents is the potentially important information that can be obtained by examining interparent agreement.

> The pattern of agreement found for parents of clinic and non-clinic children suggests that a measure of interparent agreement could be useful in child or family clinical work.... Clinically, the amount of interparent agreement could reflect a degree of interpersonal perceptivity and congruence within the family. Such a measure would be useful in identifying and beginning to treat the source and nature of the child's maladjustive behavior. (Ferguson et al., 1974, pp. 179-180)

Important information can also be obtained from an understanding of the sources of discrepancies in parents' reports and the factors related to interparent agreement. "Identifying the source of discrepancies between parent's ratings could be of clinical importance in terms of etiology, prognosis and therapy" (Thompson & McAdoo, 1973, p. 387). Potential sources of, or factors related to such discrepancies include: the child

and his or her disorder, the parents' biases, cognitions, and expectations concerning their child's behavior, and the quality of the marital relationship. These factors will be discussed in greater detail below. Although the evidence appears to be weighted more heavily in favor of the benefit, if not the necessity, of including both parent's reports, a great proportion of outcome research continues to rely primarily on maternal report.

As discussed above, resolving discrepancies between multiple sources of outcome data is a central issue in psychotherapy outcome research. When using parent data as an outcome measure, two discrepancies are possible: (1) within the parental dyad (i.e., mothers and fathers may disagree with each other), and (2) between parents and other observers (i.e., parents may disagree with reports from teachers or therapists). The former discrepancy has usually been neglected although it impacts on interpretations of the latter. It is apparent, when using the tripartite model, that it is as important to account for the possible discrepancies within each category of outcome measure as it is to consider discrepancies between categories.

Agreement between raters is usually measured by some index of reliability or agreement. Statistically, these terms are not synonymous and their differences will be discussed below, in the section on methodology. For clarity, the conventions adopted by previous authors will be used. "interparent agreement" will be used to refer to agreement within the parental dyad (i.e., between mothers and fathers). Similarly,

"interrater reliability" will refer to agreement between parents and other sources, or between other pairs of raters. The following discussion of these two types of discrepancies will be organized according to the factors that have been related to interrater reliability and interparent agreement. Research related to interrater reliability will be examined first as it is the more generally related literature. Most of the research on interrater reliability involves diagnostic uses of parent report. After discussing this general research, the literature on interparent agreement will be examined.

Interrater Reliability

Interrater reliabilities on behavior checklists have been assessed across a variety of raters (Lindholm & Touliatos, 1982). Findings have been reasonably consistent across different checklists. In a review of research on the classification of child psychopathology, Achenbach & Edelbrock (1978) concluded that interrater reliability "increased with the degree of similarity between the types of raters and between the types of situations in which the rater saw the subjects" (p. 1275). The settings, or situations, in which the ratings are made is an important variable affecting interrater reliability. It appears that changing the setting in which ratings are made confounds the effects of changes in the type of rater with changes in the child's behavior. The possibility that a child's behavior changes across settings can be argued from the concept of behavioral specificity (e.g., Novick, Rosenfeld, & Bloch, 1966). The contributions of changes in the child's behavior and changes

in raters across settings do not appear to have been isolated; subsequently, both factors need to be controlled for. In general, it appears that raters within settings (e.g., teachers-aides; mothers-fathers) agree more than raters between (across) settings (e.g., parents-teachers; parents-clinicians). Achenbach and Edelbrock (1978) report interrater reliabilities within setings to range from .50 to .78, while interrater reliabilities between settings ranged only from .25 to .51 (p. 1288).

It has been suggested that correlations within the parental dyad are higher than any other pair of raters. For example, Miller (1964), reported Q-sort correlations of .60 between parents, .37 between parents and clinicians, .35 between parents and teachers, and .24 between teachers and clinicians. This emphasis on the higher correlations within the parental dyad than between any other pair of raters is most likely due to the fact that no other within setting pairs were directly compared. Other investigators comparing raters within settings have found high correlations similar to those found between parents (e.g., Peterson, 1961; Wolf, 1981). Quay (1977) has noted this effect:

> As might be expected, the degree of agreement between raters is a function of who the raters are that are being compared and the situations in which the respective raters make their observations. Obviously, parent-parent agreement and teacher-teacher agreement is higher than parent-teacher agreement. (p. 284)

Although the distinction of whether raters are from the same or different settings is important in understanding varying levels of interrater reliability, there have been contradictory findings. Several

investigations have reported high interrater reliabilities between settings (e.g., Edelbrock & Achenbach, 1980, between mothers and clinicians) or low reliabilities within settings (e.g., Peterson, Becker, Shoemaker, Luria, & Hellmer, 1961, who found interparent correlations of only .48). Thus, interrater reliabilities appear to vary as a function of variables other than, or in addition to, the similarity of rater and setting. This will be discussed in terms of factors related to measurement and those related to informant biases.

Measurement Factors

Several factors related to the methods by which ratings are obtained affect levels of interrater reliability: the type of measure (checklist, observation), the type of rating (global vs. specific), and the type of behavior to be rated (overt vs. covert). Forehand et al. (1979) investigated the relationship among three multiple outcome measures of Parent Behavioral Training: observational data, parent recorded data and questionnaire data. Although all three measures indicated that treatment was effective there were significant relationships only within obsevational measures and within questionnaire data but not between (across) these three measures. That is, there were low correlations between data obtained from parents and independent observers. These results suggest that either no one subject demonstrated uniform changes across outcome measures or that different outcome measures assess different behaviors (Forehand et al., 1979). It should be noted that "parent" data in the Forehand et al. (1979) study included only mothers

reports and therefore no conclusions can be drawn about interparent agreement. Forehand et al.'s (1979) results do suggest that outcome measures of a given type correlate well with similar measures but not with different types of measures or with measures obtained from different sources. Other investigators have suggested that parents and independent observers can agree highly. Colletti and Harris (1977) suggested that parents are as reliable as independent observers. Interparent agreement (.90) was reported to be as high as interrater reliabilities between parents and independent observers (.87) on direct observation and recording of behaviors. It would be more accurate to refer to interobserver agreement in the Colletti and Harris (1977) study to reflect the nature of the task used. The differences in the conlusions between studies may be based on differences in the types of tasks involved (specific vs. general), the inclusion of one versus both parents, or the type of treatment under investigation (parent training, Forehand et al., 1979; vs. siblings as therapists, Colletti & Harris, 1977). In discussing factors related to measurement and calculation of interrater reliability "differences in types of rating procedures used, statistics for computing the index of reliability, and the use of general vs. situationally specific behavior samples must all be taken into account" (Burrows & Kelley, 1983, p. 42).

An additional factor affecting interrater reliability is the type of behavior to be rated. In general, higher interrater reliabilities are obtained when rating more overt, easily observable behaviors than on

ratings of covert, less observable or internal behaviors. This has been demonstrated across various type of raters. For example, researchers using the Behavior Problem Checklist have consistently found higher reliabilites for the Conduct Problem scale, measuring more overt, acting out behaviors, than for the Personality Problem scale which measures more covert, internalized behaviors (Quay, 1977). Although these differences are not always signficant, they appear to be consistent across various checklists as well as other measures. For example, higher interrater reliability for overt behaviors has also been found in the use of interviews (Graham & Rutter, 1968; Herjanic and Reich, 1982). The effect of behavior type on interrater reliability may be due to the inference required when rating covert behaviors. This inference allows for other factors to influence the obtained ratings, most importantly, informant bias.

Informant Bias

Differences in obtained ratings may reflect differences in raters' perceptions, or biases. Parents' ratings may be particularly subject to such biases due to their greater subjectivity as "significant others" (Davidson & Davidson, 1983). This subjectivity has been suggested as responsible for parents' higher ratings of outcome. "Investigators have repeatedly found that parents tend to evaluate treatment outcome more positively than either therapists...or independent observers" (Ross, 1978, p. 614). Atkeson and Forehand (1978), in reviewing studies using multiple outcome measures of Parent Behavioral Training, found that

"parent reports in terms of data concerning either specific problems or global questionnaires were associated with more positive outcomes than data collected by independent observers" (p. 456). Atkeson & Forehand (1978) further noted that parent measures agreed with each other more than with independent observers.

Sources of parental bias may be related to social desirability, parents' motivation for therapy, parents' belief that their child is maladjusted, or the type of treatment employed. Although it has been noted that parents usually report higher, more positive outcomes than other observers or raters, much of this research has been involved with studies of parent education or behavioral family therapy as treatment methodologies for dealing with children's behavioral disorders (e.g., Atkeson & Forehand, 1978; Eyberg & Johnson, 1974; Johnson & Christensen, 1975). These types of interventions might further add bias as parents are, in a sense, clients. Whether this relationship holds true for other treatment approaches is uncertain. Moreover, parent reports have usually been considered together and there is little understanding of the variability within the parental dyad.

Interparent Agreement

Interparent agreement, the second issue related to discrepancies among sources of outcome data, is a subset of interrater reliability. Most, if not all of the factors affecting interrater reliability also affect interparent agreement. Several additional factors appear to affect interparent agreement more specifically. These factors will be discussed in relation to the child, and his or her parents.

Child Factors

One of the primary factors related to the child that has been shown to affect interparent agreement is the nature of the child's behavior. As discussed above, higher interrater reliabilites are obtained for ratings of overt behavior than for covert behavior. This holds true for interparent agreement. Investigations of interparent agreement have compared agreement across scales measuring different types of behavior (e.g., Lessing & Clarke, 1982; Quay, 1977). Higher interparent agreement is typically obtained on scales measuring more overt behaviors than on those measuring covert behaviors. In terms of diagnostic categories, interparent agreement is thus expected to be higher for behavioral, or externalizing, disorders than for emotional, or internalizing, disorders.

Behavioral specificity, or situational variation in the child's behavior, may also affect interparent agreement. The effects of behavioral specificity on interparent agreement do not appear to be as well understood as its effect on interrater agreement. Burrows and Kelley (1983) reported that the familiarity of the child being rated (parents own child vs. a stranger), but not the situational specificity of the child's behavior affected interparent agreement. Novick, Rosenfeld and Bloch (1966) reported that only 6.8% of the items on which parents disagreed reflected situational variation in the child. They concluded that their findings "raise questions not about the importance of situational variation, but as to the use of parental disagreement in report as a

measure of such variation" (Novick, Rosenfeld & Bloch, 1966, p. 493). It might be hypothesized that parents see their children in more similar situations than other pairs of raters within similar settings. Burrows and Kelley (1983) reported, however, that "those parent pairs who agreed the most did not necessarily spend a large amount of time in the same kind of situations with their child" (p. 41).

Several additional factors related to the child appear to effect levels of interparent agreement. Clinic status has been shown by some investigators (e.g., Jacob et al., 1982) to be related to lower levels of parental agreement but not by others (e.g., Ferguson et al., 1974). Ferguson et al. (1974) suggested that mothers and fathers of clinic and nonclinic children "may be sensitive to different behavioral dimensions" (p.179). Similarly, the childs sex and age appear to be related to levels of agreement, although perhaps not directly. Thompson and McAdoo (1973) report interparent agreement (correlations) of between .40 and .72 for boys, and between .20 and .78 for girls across 7 scales of behavior ranging from somaticization to aggression. Although interparent agreement appears quite variable across different behavioral dimensions, Thompson and McAdoo (1973) reported no significant differences between parents for boys and only one significant difference for girls, such that fathers rated significantly more sociability for girls than did mothers. Achenbach and Edelbrock (1979) reported mean levels of interparent agreement, averaged across scales, to be .79 for boys, .63 for 6- to 11-year-old girls, and .54 for 12- to 16-year-old girls.

Again, they reported no significant differences between parents but did comment on a trend for mothers to report more problems for 6- to 11-year-old girls than fathers.

Ferguson et al. (1974) found no main effects for the child'd sex, age and adjustment (clinic status) but did find interactions among these variables to be related to levels of interparent agreement. Specifically, "the parents of 5- to 7-year-old nonclinic males showed the greatest amount of agreement while the parents of 5- to 7-year-old clinic males showed the lowest" (Ferguson et al., 1974, p. 180). They concluded that "the interparent agreement index may be most useful as a predictor of adjustment among younger boys" (Ferguson et al., 1974, p. 180).

To summarize, it appears that although the child's age and sex do not necessarily exert direct influences on interparent agreement, they may interact and may, in combination with other factors, affect levels of interparent agreement. Moreover, there is general consensus on the need to control for age and sex when studying children. Additionally, the severity and chronicty of the child's disorder have been suggested as important variables to control. Although severity and chronicity have not been examined in relation to interparent agreement, they are potentially important factors.

The present study will examine interparent agreement as a function of: (a) the type of disorder manifested by the child, as rated by the clinic staff (behavioral problem vs. emotional problem), and as based on

the presenting problem (overt vs. covert); (b) whether the child's disorder is manifested only at home or not, as determined by the nature of the presenting problem and the type of referral source; (c) the severity of the child's disorder, as calculated from clinic staff ratings; (d) the chronicity (duration) of the child's disorder, as reported by his or her parents; (e) whether or not the child had been previously evaluated, as reported by the parents; (f) the child's age; and (g) the child's sex. These variables will be referred to collectively as "child factors."

Parent Factors

Several factors related to the child's parents and their perceptions and labelling of their children's behavior may affect levels of interparent agreement. Interparent agreement may vary as a function of parent's interpersonal perceptions (Ferguson et al., 1974; Novick, Rosenfeld & Bloch, 1966) or labelling or defining of behaviors (Ferguson et al., 1974; Thompson & Bernal, 1982). Moreover, parents perceptions and labelling of their children's behavior may be related to (a) parental pathology (Christensen, Phillips, Glasgow, & Johnson, 1983; Forehand, Wells, McMahon, Griest, & Rogers, 1982; Griest & Wells, 1983), (b) the quality of the marital relationship (Christensen, Phillips, Glasgow and Johnson, 1983; Emery & O'Leary, 1984; Porter & O'Leary, 1980; Thompson & Bernal, 1982), and (c) various parental cognitive factors (Christensen et al., 1983; Rickard, Graziano, & Forehand, 1984; Thompson & Bernal, 1982). These three sets of factors have not been directly exam-

ined in relation to levels of interparent agreement, however. That is, interparent agreement appears to be a function of parents' perceptions of their children which, in turn, may be a function of the factors described above.

Forehand et al. (1982) reported that "maternal depression was the best predictor of maternal perception of children" (p. 145). Similarly, Griest and Wells (1983) found that maternal report of childhood behavior varies as a function of maternal depression and anxiety. Both of these studies (Forehand et al., 1982; Griest & Wells, 1983) relied soley on maternal reports and interparent agreement was not examined. Similarly, studies examining the quality of the marital relationship to child behavior problems have often relied on data obtained soley from mothers (e.g., Emery & O'Leary, 1984; Porter & O'Leary, 1980).

Christensen et al. (1983) examined the intercorrelations of four family variables: (a) marital disturbance, (b) parental psychopathology, (c) interactional dysfunction, and (c) parental cognitions. They examined the relationships between "measures of marital discord, parental psychopathology, and three parental cognitive factors: knowledge of behavioral principles, tolerance for child deviancy, and expectations regarding the child's behavior" (Christensen et al., 1983, p. 153). Additionally, recordings of random family interactions were made at home to assess interactional dysfunction. Unfortunately, although ratings were obtained from both mothers and fathers, no comparison was made between them. "Marital discord was the only measure associated with

parental perceptions of child problems" (Christensen et al., 1983, p. 153). Data from other studies suggest that the effects of the quality of the marital relationship on interparent agreement appear to be equivocal, however. Thompson and Bernal (1982) reported that marital distress, parent behaviors toward the child, and parental distress about the child's deviant behavior were neglibly related to parental labelling. These contradictory findings may be due to differences in the samples, methodologies and statistics utilized. Robinson and Anderson (1983) have found, however, that when social desirability is controlled for, the marital-child adjustment relationship was non-significant. Horton (1984) suggests that "parent reported marital satisfaction must be interpreted cautiously" (p. 275). The effects of the marital relationship on parental perceptions of the child and the effects of treatment do not, then, appear to be fully understood at the present time. Similarly, the effect of parental pathology and cognitive factors on parent's perception of their child's behavior needs further examination. Moreover, the relationship of these factors to interparent agreement needs to be elucidated, as it may be appropriate to combine parent's reports in some instances while it might be misleading to do so in others.

One additional factor potentially related to interparent agreement remains to be explored. The research directly examining variables related to interparent agreement has obtained parent ratings on checklists utilized for diagnostic purposes (e.g., Ferguson et al., 1974;

Jacob et al., 1982). Although changes in scores on parent completed checklists have been utilized as an outcome measure, factors affecting interparent agreement on "post", or outcome, checklist scores has not been evaluated. The potential effect of using parent ratings to gather diagnostic information versus the assessment of treatment outcome on interparent agreement may be clinically useful information and should be examined. Such differential effects, should they exist, might provide information about changes in parents as well as in their children. For example, levels of interparent agreement may increase or decrease during the course of the child's treatment, and may affect the direction of outcome ratings. The present study will compare factors affecting interparent agreement on pre- and post-treatment administered checklists.

In summary, a number of factors related to the child's parents, their individual mental health, the quality of their marital relationship, and their cognitions and expectations regarding their children's behavior may affect levels of interparent agreement on checklist ratings. These variables have not all been directly examined in relation to interparent agreement but have been shown to be related to parent's perceptions of their children's behavior. Moreover, these variables may be further indicative of the parent's relationships to each other and to their children and may provide important information for the diagnosis and treatment of children's disorders. The relationship of these variables to interparent agreement needs to be examined.

The present study will examine interparent agreement as a function of: (a) parental pathology, as measured by staff ratings of the need for individual treatment and mother's treatment history, (b) the quality of the marital relationship, as measured by staff ratings of the need for marital treatment, (c) mothers' and fathers' level of education, (d) parental cognitions and expectations about their child's problems as measured by openended questions on a checklist, the parents' level of education, and staff ratings of the quality of child rearing techniques and parents' level of motivation. These factors will be referred to collectively as "parent factors." Additionally, the relationship of interparent agreement to the type of treatment received, the type of referral source, and pre- versus post-treatment checklist measures will be examined. These will be referred to collectively as "treatment factors."

In conclusion, interparent agreement is generally moderate to moderately high. However, caution should be exercised before concluding that parents' reports are similar enough to be combined. For example, Achenbach and Edelbrock (1978) reported interparent correlation coefficients to be within a .50 to .76 range. Although, these coefficients are all significantly different from zero, they provide no information about the variablility of levels of interparent agreement for individual parental dyads. As discussed above, it is important to account for such individual variability. Jacob et al. (1982) report that the "average level of agreement is moderate, while the variability in levels of

agreement is considerable" (p.603). Additionally, interparent agreement has been been found to vary as a function of the child's clinic status, clinical diagnosis, sex, age, and behavior to be rated; and may be related to the quality of the marital relationship; and other factors related to parental perceptions and expectations. Jacob et al. (1982) suggest the following areas for subsequent research efforts: (1) the use of larger, more representative samples;, (2) the systematic study of the relationships between parent agreement and (a) interparent consistency in interactions with the child, (b) the nature and severity of childhood disturbance, and (c) the level of distress/dysfunction in the marital dyad; and (3) the separate assessment of each parent's perceptions of his/her child's behavior. The present study is intended to address some of these recommendations.

Methodological Issues

Agreement vs. Reliability

In the literature the terms reliability and agreement are often used interchangeably although distinctions can be made on the bases of when and how they should be used, how they are calculated, and the information they provide. Various types of reliability and agreement can be utilized depending on the level of measurement in a given study. The present discussion will focus on the use of scaled data at either the ordinal or interval levels of measurement.

Interrater agreement usually "represents the extent to which different judges tend to make exactly the same judgements about the rated

subject" (Tinsley & Weiss, 1975, p. 359). Agreement between raters on behavior checklists is usually calculated by dividing the number of agreements on individual items by the sum of agreements and disagreements (A/A+D). This results in the "Percentage Agreement" (PA) between raters, one of the most frequently used indices of agreement. PA as an index of agreement has several disadvantages; it includes chance agreement, is insensitive to degrees of agreement (agreement is all or none), and is usually not associated with a significance test (APA, 1982). Additionally, it is affected by the frequency of endorsement, that is the number of items endorsed, because it includes agreement on the non occurrence of items. Thus, the level of agreement is affected by the number of items endorsed by the raters. If raters endorse a small number of items, their level of agreement will be inflated by the high number of agreements on items not endorsed (left blank by both raters). Thus differences between the findings of various studies may be due to differences in methods of calculating agreement. In conclusion, "the alternative reliability indices available to represent interparent agreement yield different patterns of results, largely as a function of their statistical properties" (Jacob et al., 1982, p.606)

Jacob et al. (1982) suggest two alternative means of calculating interparent agreement: the "Effective Percentage Agreement" (EPA, the PA for only those instances where either or both raters check an item); and the "Absolute Difference" (AD, the absolute value of the difference between Parents' total scores). The EPA avoids the frequency of

endorsement problem by only counting agreement on items that are endorsed by at least one of the raters. Similarly, "the proposed Absolute Difference score appears to be an appropriate and sensitive measure of parental agreement" (Jacob et al., 1982, p. 606) that is not spuriously effected by frequency of endorsement. The AD score is useful for "investigations of relationships between agreement and other variables" (Jacob et al., 1982, p. 594) but does not provide any information about agreement on individual items.

Interrater reliability, on the other hand, is related to the relative ordering of subjects by raters or "the degree to which the ratings of different judges are proportional when expressed as deviations from their own means" (Tinsley & Weiss, 1975, p. 359). While interrater agreement provides no information about the variability among subjects, interrater reliability is very much concerned with the variability among subjects (APA, 1982). Interrater reliability is usually expressed "in terms of correlational or analysis of variance indices that portray whether the average difference between subjects is large relative to the degree of disagreement between judges" (APA, 1982, p. 96). One drawback with correlational indices is that the significance is usually reported in terms of the difference of the coefficient from zero; statistically significant correlations may account for mimimal amounts of variance.

As a final distinction, it is important to note that reliability and agreement are not equivalent or directly related indices.

> High reliability is no indication that the raters agree in an absolute sense on the degree to which the ratees possess the characteristic being judged...On the other hand, low reliability does not

> necessarily indicate that the raters are in disagreement. (Tinsley & Weiss, 1975, p. 360)

Choosing the appropriate index depends on the nature of one's data and the questions that are to be answered. If relative orderings among raters are of primary concern, the interrater reliability index is a satisfactory index of the ratings. Reliabilty indices do not, however, make evident the differences among judges in their ratings. Therefore, whenever the absolute value of the ratings as defined by points on a scale are of importance, interrater agreement should be utilized. When assessing interparent differences, agreement indices appear to be more approprite. When calculating interrater differences, however, indices of reliability may be more appropriate. In summary, indicies of reliability and agreement provide different information that may be more or less appropriate to a given study. Moreover, different methods of calculating these indices yield different results based on their statistical properties. An awareness of these methodological issues is important when interpreting the results of a study and when comparing different studies. Finally, in order to fully understand the relationships among raters, indices of both reliability and agreement should be utilized when possible.

Measuring Change

The issue of how, and whether, to measure change is still being debated (e.g., Cronbach & Furby, 1970). This issue is complicated by the fact that little has been written about the methodology for measur-

ing change that is comprehensible to clinicians or aimed directly at psychotherapy (Mintz et al., 1979). Currently, there are several resolutions to the debate. Among the current methodologies for assessing change in use are: (a) raw change, or gain, scores, (b) proportional gain scores, (c) Analysis of Covariance (AVCOVA), and (d) residual gain scores. Additionally, there are two measurement strategies that avoid using change as the criterion. These are: (a) final adjustment status, and (b) direct ratings of benefit. All of these methodologies have advantages and disadvantages. The selection of the most appropriate method for assessing change is dependent upon the needs, assumptions, and methods of a given study. Further analysis of these methods and their appropriateness is beyond scope of the present study. The reader is referred to APA (1982) and Mintz et al., (1979) for a more detailed discussion of the current methodologies for assessing change.

Summary

In summary, the purpose of the present study is to examine methodological issues in the use of parent report data to assess the outcome of child psychotherapy. More specifically it is intended to examine: (1) the effects of using maternal and paternal reports separately and in combination with each other, (2) the factors affecting mothers' and fathers' perceptions of their children's behavior as measured on a behavior checklist, both before and after treatment, (3) the level of agreement between parents regarding their clinic referred child, and (4) the possible factors related to the child and his or her disorder, the

child's parents, and the type of treatment approach employed that might affect interparent agreement before and after treatment. Finally, this study will provide further information about the clinical application of the primary measure utilized, the Washington Symptom Checklist (Wimberger and Gregory, 1968).

CHAPTER III

METHOD

Setting

The data for the present study are based on archival data obtained from a multidisciplinary outpatient child guidance clinic affiliated with a United States Army hospital in Germany. The clinic served a population of 55,000 military personnel and their dependent families. Civilian employees of the United States government were also eligible for treatment. The data were collected during a three year period in the early 1970's.

At the time of data collection, the clinic conducted a traditional intake and treatment procedure. The intake portion of this procedure involved an application by the parents to the center, a social history intake interview conducted with both parents, an individual interview/diagnostic session with the identified patient (the child), a teacher's report from the child's school, and in some cases a school visit and/or classroom observation by the clinic staff. After the intake procedures were completed, staffings were held at the clinic to review the case and determine treatment disposition. Staff Psychologists, Psychiatrists, and Social Workers, as well as para-professional members of the clinic staff were present at the staffings. Treatment assignments varied, as

did the length of involvement in treatment. Treatment approaches utilized included individual and group therapy for the identified patient, couple and group therapy for the child's parents, and, in some cases, conjoint treatment for the child and his parents.

Subjects

Data were available on 235 families, most of whom were lower to middle class in socioeconomic status. Sixty percent of the sample was from the families of enlisted men, 32% of the sample was from the families of officers, and 8% of the sample was from civilian families. Fourty three cases were dropped from the total sample because the identified patient was over 12 years old. Three additional cases were dropped because the child's age was not available. These exclusions accounted for 19.6% of the total sample. Of the remaining 189 cases in which the child was 12 years old or younger, 27 additional cases were dropped. Of these 27, 15 cases were excluded because they had refused intake or treatment, or had dropped out of treatment; 10 cases were excluded because parent report data were missing; one case was dropped because the family had moved during treatment; and one case was dropped because the child was diagnosed as schizophrenic.

The remaining sample of 162 cases consisted of the families of 120 male children (74%) and 42 female children (26%). This ratio of males to females is generally consistent with child guidance center populations. The mean age for the total sample was 8 years (males $\underline{M}$ = 7 years, 11 months; females $\underline{M}$ = 8 years, 4 months). The final sample

thus consisted of the families of all children receiving outpatient treatment service with the following exceptions: Psychosis, mental retardation, known or suspected organic brain damage, all adolescents (12 years or older), and families offered only crisis intervention or emergency services. The children studied include those considered to have character disorders, behavior disorders, and neurotic disorders although no formal diagnostic classifications were available. The mean duration reported for presenting problems was 28 months (SD = 18 months). 24 months was the modal length of duration. Of the 162 cases in the final sample, 38 were on a waiting list control and 124 were in treatment.

Measures

Staff Ratings. Two measures were collected from clinic staff; one was diagnostic in nature and the second was an outcome measure assessing the effects of treatment. The first (diagnostic) clinic measure consisted of ratings completed by all staff in attendance at the case staffing during which the treatment disposition was decided. After all diagnostic information was presented (e.g., results of social history interview, diagnostic session with the child, etc.,) each staff member independently rated the case on six dimensions. On each dimension 7 point ratings were possible, with higher scores reflecting more severe problems or less motivation. The six dimensions rated were as follows: (1) the presence of Minimal Brain Dysfunction (MBD) symptoms. (2) the

degree of behavioral management problem, (3) the degree of emotional problem, (4) the ineffectiveness of each parent's child rearing techniques, (5) the need for marital/individual therapy, and (6) each parent's lack of motivation to seek help. An average score for each dimension was calculated across all raters. A copy of this case evaluation form is in Appendix A.

The outcome measure collected from clinic staff consisted of therapist ratings of improvement on a 7 point scale (-3 = very much worse; 0 = no change; +3 = greatly improved).

Parent Ratings. Parental ratings were obtained separately for mothers and fathers on the Washington Symptom Checklist (WSCL) (Wimberger and Gregory, 1968). The WSCL consists of 66 behavior statements to be rated on a four point scale as to frequency of occurrence (never, seldom, frequently, or very often, scored 0, 1, 2, and 3, respectively). Additionally, the WSCL includes nine questions designed to assess parents' general attitude and expectations about their child's problems, parent agreement about the problems and degree of parental pathology and motivation. For the present study, only the first five of these nine questions are utilized. These five questions are answered on a three point scale (yes, undecided, no). Item scores for the 66 behavior statements are unavailable. Total scores for positive and negative behaviors (items) will be utilized separately in analyzing the data.

Test-retest reliabilities calculated over a 30 day interval were .84 for the parents of clinic children (N = 66), and .87 for parents of

nonclinic children (N = 74; Wimberger & Gregory, 1968). Evidence for the validity of the WSCL is reported in terms of high agreement between parent and therapist ratings (N = 22, $p<.001$; Wimberger & Gregory, 1968). Wimberger and Millar (1968), concluded that "the WSCL appears to be a valid instrument in that it measures factors which are verifiable by other reliable sources. Furthermore, it contains items pertaining to the general emotional well being of the child which serve as good indices distinguishing clearly the control and clinic groups" (1968, p. 182). Humphreys and Ciminero (1979) suggested, however, that "the data on the reliability and validity of the WSCL are inconclusive" (p. 58). Additional concerns about the WSCL include the lack of norms or an established cutoff score, and the lack of separate factor scores to distinguish different types of disorders (e.g., internalizing vs. externalizing).

CHAPTER IV

RESULTS

Means and Correlations

As a reliability check on the WSCL, the correlation between parents' checklist scores at the time of intake and immediately prior to the initiation of treatment was calculated for 38 subjects on a waiting list control. The mean length of time between these measures was 13.6 weeks (SD = 4.96). Twelve weeks was the modal interval between administrations. The reliability coefficients for mothers' ratings of positive and negative items were .50 and .84 respectively. For fathers', these coefficients were .71 and .84, respectively. All coefficients were significant at or beyond the .001 level. The reliability for negative items was higher than positive items for both mothers and fathers. These findings are consistent with the coefficient of .84 reported by Wimberger and Gregory (1968) for 66 parents of 40 clinic children over a 30 day period. Wimberger and Gregory (1968) did not separate positive and negative items and their coefficients represent reliability for total item scores. However, it should be noted that some intervention, in the form of intake interviews and procedures, occurred between the administrations of the checklist for waiting list subjects in the present study. Such procedures typically produce changes in data collected

from waiting list controls in most psychotherapy research. Therefore, the current effort to assess reliability does not represent a true test-retest procedure and the resultant data are probably an underestimate of the stability of checklist scores over time.

Pretreatment

Table 1 presents the means and standard deviations for mothers' and fathers' pretreatment checklist scores for daughters, sons and the total sample. Mean differences between mothers' and fathers' scores were not significant for daughters, sons, or the total sample on either positive or negative items. Similarly, neither parent rated sons and daughters significantly different on either positive or negative items. Thus, the child's sex did not significantly affect parents' ratings. In order to gain a better understanding of the relationship between mothers' and fathers' scores, correlation coefficients, mean Difference (D), and mean Absolute Difference (AD) scores were also examined.

Table 2 presents the means and standard deviations for pretreatment D and AD scores for daughters, sons and the total sample. D scores were calculated by subtracting father's score from mother's score for each parental dyad. AD scores are the absolute value of D scores. D scores thus provide information about which parent's score is higher, whereas AD scores provide a general index of interparent agreement regardless of which parent's score is higher. Mean D scores for sons and daughters were not significantly different for positive or negative items. Similarly, mean AD scores for sons and daughters were not sig-

TABLE 1

Mothers' and Fathers' Pretreatment Checklist Scores

		Mother		Father	
		Positive Items	Negative Items	Positive Items	Negative Items
Daughters	*M*	9.36	81.50	9.10	81.88
	SD	3.12	17.59	2.71	14.65
Sons	*M*	8.77	86.37	8.62	83.56
	SD	2.84	16.62	2.45	18.31
Total	*M*	8.92	85.11	8.74	83.12
	SD	2.92	16.96	2.52	17.40

Note. Daughters (n = 42); Sons (n = 120); Total (N = 162).

nificantly different for positive or negative items. This suggests that the child's sex is not a significant determinant of levels of interparent agreement.

Two-tailed *t*-tests were used to determine if the obtained D and AD scores were significantly different from the expected population mean of zero (reflecting complete agreement between parents). Only one mean D score was negative, indicating that fathers tended to rate daughters more negatively than did mothers on negative items. However, this D score was not significantly different from zero. One mean D score was significantly different from zero, indicating that mothers rated sons more negatively than fathers on negative items ($t(119) = 2.1$, $p<.05$). All mean AD scores were significantly different from zero indicating that there was a significant amount of non-agreement between parents (All *t*'s > 7.3, $p<.001$; *df* = (41) daughters, (119) sons). This suggests that mean AD scores of zero are not to be expected.

Table 3 presents the correlation coefficients between mothers' and fathers' pretreatment checklist scores for daughters, sons and the total sample. The range for interparent correlations, .32 to .65, is somewhat lower than might be anticipated. This appears to be due to two factors: (a) the distinction made between positive and negative items on the WSCL, and (b) the lack of separate scale scores on the WSCL. The distinction between positive and negative behaviors or items is not ususally made in the literature because most checklists focus on negative or problematic behaviors. In general, interparent correlations were lower

TABLE 2

Pretreatment Difference and Absolute Difference Scores

		Positive Items	Negative Items
		Difference Scores	
Daughters	M	.26	-.38
	SD	2.70	15.83
Sons	M	.15	2.81*
	SD	3.11	14.62
Total	M	.18	1.98
	SD	2.99	14.96
		Absolute Difference Scores	
Daughters	M	2.02**	12.48**
	SD	1.77	9.55
Sons	M	2.32**	11.21**
	SD	2.07	9.76
Total	M	2.24**	11.54**
	SD	1.99	9.69

Note. Daughters (n = 42); Sons (n = 120); Total (N = 162).
*p<.05. **p<.001.

for positive items than for negative items. Nevertheless, the overall correlation for negative items, .62, is within the range of .50 to .76 reported by Achenbach and Edelbrock (1978) for parent ratings. The lack of separate subscales on the WSCL necessitates that interparent correlations be based on heterogeneous rather than homogeneous groups of items. Homogeneous groups of items are more likely to yield higher correlation coefficients. A more specific analysis of interparent correlations would be possible if the WSCL had separate scales for different types of disorders.

It should be noted that the use of correlation coefficients (a measure of reliability) and AD scores (a measure of agreement) may provide different information. For example, the correlation coefficent between parents on positive items for males is only .32, suggesting poor reliability, whereas the low mean AD score ($\underline{M}$ = 2.32) suggests relatively high agreement. The opposite appears to hold true for negative items for males, where the correlation coefficient suggests moderate reliability ($\underline{r}$ = .65). but the mean AD score ($\underline{M}$ = 11.21) suggests low agreement. Conclusions based on these comparisons are tentative, however, as more information is needed about levels of AD scores that constitute significant levels of disagreement. Moreover, levels of reliability and agreement for positive and negative items on the WSCL are not directly comparable as there are fewer items on the positive dimension ($\underline{n}$ = 6) than on the negative dimension ($\underline{n}$ = 61). Additionally, the use of means and D scores may also provide different information.

For example, the mean D score for parents' ratings of males on negative items is significantly different from zero, but the difference between mothers' and fathers' mean scores is not significant. It can be concluded that interpretations based on mean scores, correlation coefficients, mean D, and mean AD scores may provide different results and that caution should be exercised when comparing findings based on different indices.

Posttreatment

Table 4 shows pre- and post-treatment means for positive and negative items for mothers, fathers and an average of their scores, obtained by dividing their sum in half, for the total sample. In general, mothers' and fathers' scores increased for positive items and decreased for negative items suggesting that therapeutic changes had occurred for the group as a whole. However, more detailed analyses of these changes and rival explanations for the change, such as regression toward the mean, are beyond the scope of the present study and should be conducted before using these data to support treatment effects.

Two analyses were conducted to examine how change over time is reflected by the use of mothers' and fathers' reports separately and in combination. The first of these involves the calculation of an effect size (Cohen, 1977). The second involves the calculation of proportional gain scores for parents' ratings and then the correlation of these gain scores with therapists' ratings of improvement. Table 4 shows the effect size coefficients for negative and positive items calculated from

TABLE 3

Pretreatment Interparent Correlation Coefficients

	r mother.father	
	Positive Items	Negative Items
Daughters	.58**	.53**
Sons	.32**	.65**
Total	.40**	.62**

*p<.05. **p<.01.

mothers' and fathers' reports separately, and from an average of mothers' and fathers' scores. The effect size is calculated by dividing the difference between post- and pre-treatment scores by the pretreatment standard deviation. It can be seen from Table 4 that the effect of averaging mothers' and fathers' reports is to obtain a mean and standard deviation between mothers' and fathers' individual means and standard deviations. The reduction in standard deviation appears to be most critical as it creates an effect size score greater than would be obtained using individual parent scores in isolation. These data suggets that effect size scores based on an average of parent ratings inflate estimates of outcome over those based on separate parent scores.

Secondly, to obtain a measure of change based upon parents' reports, a proportional gain score was calculated. A proportional gain score is the difference between pre- and post-treatment scores divided by the pretreatment score. The resulting gain score is thus proportional to the intial level and effectively controls for differences in initial level. Proportional gain scores were calculated for mothers' and fathers' reports separately, for an average parental score, and for the AD score between parents. Table 5 shows the correlation between these proportional gain scores and therapists' improvement ratings for children in treatment conditions with and without parental involvement. None of the correlations are significant for the child only treatment group (without parents), but there are significant correlations for treatment that included parents. This suggests that the degree of cor-

TABLE 4

Mean Checklist Scores and Effect Size for Treatment Sample

		Mother		Father		Parental Average	
		Positive Items	Negative Items	Positive Items	Negative Items	Positive Items	Negative Items
Pre	M	8.93	85.91	8.67	84.36	8.80	85.13
	SD	2.89	16.69	2.32	16.74	2.13	15.08
Post	M	10.03	71.41	9.46	71.12	9.74	71.26
	SD	2.94	15.16	2.53	15.89	2.32	13.95
Effect Size		.38	-.87	.34	-.79	.44	-.92

Note. n = 124.
Levels of Effect Size: .2= small;.5= medium; .8= large.

respondence between parents' and therapists' reports of child treatment outcome might be a function of whether or not parents are also involved in treatment. Moreover, relying on an average of mothers' and fathers' scores might, again, be misleading. The significant correlation between the average proportional gain score for parents and the therapists' outcome rating appears to be due primarily to the influence of the correlation between fathers' proportional gain score and the therapists' rating. The correlation bewteen mothers' and therapists' scores is not significant. This difference between mothers and fathers would be obscured by relying on the average of parents' scores. The negative value of the significant correlations indicates that decreases in fathers' and the average parent ratings are associated with therapists' ratings of improvement. This relationship is in the expected direction.

Multiple Regression Analyses

Pretreatment

Using mothers' and fathers' scores as separate criterion variables, stepwise regression analyses were conducted to determine which of the predictor variables (child factors and parent factors) accounted for significant amounts of variance in each parent's perception of his or her child's behavior. Stepwise regression analyses were also done using D and AD scores as criterion variables to determine what factors affect interparent agreement. Listwise deletion of missing data was used for all regression analyses.

TABLE 5

Correlations Between Parents' and Therapists' Ratings

Proportional Gain Scores	Treatment Condition: Without Parents	Treatment Condition: With Parents
Mother		
Negative	ns	ns
Positive	ns	ns
Father		
Negative	ns	-.41**
Positive	ns	ns
Parental Average		
Negative	ns	-.31**
Positive	ns	ns
Absolute Difference		
Negative	ns	ns
Positive	ns	ns

Note. ns = not significant.
**$p < .01$.

Significant first order correlation coefficients between mothers' and fathers' checklist scores and predictor variables used in the multiple regression analyses are in Appendix C. Significant first order correlation coefficients between D and AD scores and predictor variables used in the multiple regression analyses are in Appendix D. Nonsignificant correlations for variables that entered in regression analyses are also included in these appendices.

For the regression analyses, variables are listed in the tables and discussed in the text in the order of the amount of variance for which they accounted (most to least). Table 6 shows the variables that entered on the stepwise regressions for mothers' pretreatment checklist scores. Three variables emerged as significant predicitors of mothers' positive ratings. For positive items, mothers' ratings were higher if they had completed a high school education, if the child's problems were rated by the clinic staff as being more emotional than behavioral in nature, and if the child was younger. One variable emerged as a significant predictor of mothers' negative ratings. For negative items, mothers' ratings were higher as a function of the increasing severity of the child's problem. The rating of severity was based upon clinic staff ratings of the degree of emotional and behavioral disturbance.

Table 7 shows the variables that entered on the stepwise regression for fathers' pretreatment checklist scores. Three variables emerged as significant predictors of fathers' positive ratings. For positive items, fathers' ratings were higher if the child's problems

TABLE 6

Pretreatment Regressions on Mothers' Checklist Scores

Source	Beta	F	Semipartial r^2	Multiple R	R^2
Positive Items					
High School	.31	12.47**	.10	.32	.10
Emot/Beh Prob	-.24	7.32**	.04	.38	.14
Age	-.18	4.03*	.03	.42	.18
Negative Items					
Severity	.29	10.07**	.08	.29	.08

Note. n = 110.
*p<.05. **p<.01.

were rated by the staff as more emotional than behavioral, if fathers reported that their spouse did not agree with them about their child's problems, or if they reported that they were not embarassed by their child's problems. Four variables emerged as significant predictors of fathers' negative ratings. For negative items, fathers' gave higher ratings if they felt their child had an emotional problem, if they reported being embarassed by their child's problem, if the staff rated the problem as more behavioral than emotional, and if the child was older.

Table 8 shows the variables that entered on the stepwise regression for pretreatment D and AD scores. The factors related to D scores will be discussed first. One variable emerged as a significant predictor of the difference between parents' ratings on positive items. For positive items, mothers' ratings exceeded fathers' ratings if the mothers' had completed high school. Five variables emerged as significant predictors of the difference between parents' ratings on negative items. For negative items, mothers' ratings exceeded fathers' ratings if the mothers' reported that they felt their child had an emotional problem; fathers' ratings exceeded mothers' ratings if the fathers' reported being embarassed by their child's problem, if they had completed college, if the child was older, and if they reported that they felt that their child had an emotional problem.

Absolute Difference scores, a measure of interparent agreement, were affected by a different set of factors. One variable emerged as a

TABLE 7

Pretreatment Regressions on Fathers' Checklist Scores

Source	Beta	F	Semipartial r^2	Multiple R	R^2
Positive Items					
Question C	-.28	9.98**	.07	.26	.07
Emot/Beh Prob	.23	6.73*	.05	.34	.12
Question B	-.19	4.79*	.04	.40	.16
Negative Items					
Question A	.34	18.54**	.14	.38	.14
Emot/Beh Prob	.33	17.38**	.08	.46	.22
Age	.29	13.34**	.09	.55	.31
Question B	.19	5.76*	.03	.58	.34

Note. n = 115. Questions: A "Do you think your child has an emotional problem?"; B "Does it embarass you?"; C "Does your spouse agree?"
*p<.05. **p<.01.

significant predictor of interparent agreement on positive items. For positive items, parents' responses to the question "Does your spouse agree that your child has a problem?" accounted for a significant amount of variance in the AD score. For this analysis, parents' responses to this question were coded contingent upon their spouse's response. The way this variable was coded indicated that, for positive items, interparent agreement was higher (AD was lower), if both parents reported that their spouse agreed with them about their child's problem. Agreement was lowest if parents responded differently about whether their spouse agreed with them. If both parents reported that their spouse disagreed with them, agreement was between that of the other two conditions. Thus, it may be more important that parents are in agreement about the status of their marital relationship as it relates to parenting, be it one of agreement or disagreement, than the actual level of agreement. Three variables emerged as significant predictors of interparent agreement on negative items. For negative items, interparent agreement was higher (AD was lower), if the level of marital pathology, as measured by the staff rating of the need for marital treatment, was lower; if the child's presenting problems were rated as more overt than covert; and if the child had not been previously evaluated.

Posttreatment

To determine what factors contributed to posttreatment scores separate multiple regressions were again conducted for mothers', fathers' checklist scores and for D and AD scores. In order to control

TABLE 8

Pretreatment Regressions on D and AD Scores

Source	Beta	F	Semipartial r^2	Multiple R	R^2
Difference Scores					
Positive Items					
High School-M	.28	9.15**	.08	.28	.08
Negative Items					
Question B-F	-.29	11.70**	.09	.29	.09
College-F	-.26	9.25**	.05	.37	.13
Age	-.23	6.96**	.04	.42	.18
Question A-M	.27	9.21**	.04	.47	.22
Question A-F	-.21	5.72*	.04	.51	.26
Absolute Difference Scores					
Positive Items					
Spouse Agree?	.27	4.20*	.07	.27	.07
Negative Items					
Marital	.39	19.90**	.15	.39	.15
Overt/Covert	-.25	8.43**	.06	.45	.20
Prev.Eval.-C	.17	4.04*	.03	.48	.23

Note. n = 109. Questions: A "Do you think your child has an emotional problem?"; B "Does it embarass you?"; C "Does your spouse agree?" Abbreviations: M-Mother; F-Father; C-Child. *p<.05. **p<.01.

for the effects of initial level on posttreatment level the corresponding pretreatment score was entered into each posttreatment equation first to remove its influence from the model. That is, mothers' pretreatment positive scores were entered as the first variable in the regression equation to predict mothers' posttreatment positive score. A similar procedure was followed for each criterion variable. Stepwise regressions were then conducted to determine the amount of variance accounted for by a subset of treatment related predictor variables: whether or not the parents were involved as the referral source, whether or not the parents were involved in treatment, whether the problem occurred at home or only at school, and the number of treatment sessions. Overall, little additional variance was accounted for beyond that of the pretreatment scores. However, several variables did account for significant, although minimal, amounts of variance in posttreatment checklist scores.

For positive items, referral source accounted for 3.3% of the variance in fathers' posttreatment checklist scores after pretreatment scores were controlled. The direction of the correlation suggests that, for fathers, self referral is related to lower posttreatment checklist scores on positive items than if the parents were not a referral source. For positive items, the number of parent treatment sessions accounted for 2.6% of the variance in mothers' posttreatment checklist scores after pretreatment score was controlled. This suggests that, for mothers, attending a greater number of parent treatment sessions was associ-

ated with a decrease in the number of positive checklist items reported posttreatment. None of the variables accounted for a significant amount of variance for posttreatment D scores on positive items. Referral source accounted for 4.3% of the variance in posttreatment AD scores after pretreatment scores had been controlled. Posttreatment AD scores were higher if the parents were self referred or were at least one of the referral sources.

For negative items, none of the variables accounted for a significant amount of variance in posttreatment scores for fathers', D, or AD scores. Treatment type accounted for 1.8% of the variance in posttreatment checklist scores for negative items for mothers. This suggests that, for mothers, greater decreases in the number of negative items reported were associated with treatment modalities that did involve the parents.

In summary, although minimal amounts of variance in posttreatment scores were accounted for beyond that of pretreatment scores, several treatment factors did account for statistically significant amounts of variance. This suggests that factors related to the source of referral, and the type and length of treatment warrant further study as possible influences of parents' perceptions and of interparent agreement on posttreatment measures.

CHAPTER V

DISCUSSION

The findings of the present study support the contention that information should be obtained from both parents about their children's behavior. Moreover, mothers' and fathers' perceptions appear to be influenced by different factors and provide potentially unique views of the child. The present findings support Jacob et al.'s (1982) conclusion that mothers' and fathers' reports should be obtained and analyzed separately. Furthermore, such a procedure appears to be important both for diagnostic purposes and for assessing the outcome of treatement.

Means and Correlations

The present findings support the contention that important information may be lost or obscured if only one parent is relied upon for information or if parents' reports are averaged together. Although differences between the means of mothers' and fathers' reports are not significant and mean D and AD scores are moderate, relying on such data to contend that parents' reports are similar enough to be interchangeable does not appear appropriate. It must be noted that the standard deviations for these distributions are relatively large, suggesting a high degree of variability for individual parental dyads. These data are consistent with the findings of Jacob et al. (1982) that average levels

of interparent agreement are moderate, but variability is high. In effect, this high degree of variability is obscured if only group means are analyzed. Secondly, the findings of the present study support the contention that different measures yield different conclusions about levels of parental agreement. Correlation coefficients, means, D scores and AD scores all provide somewhat different information. Therefore, the findings of previous studies using different indices may not be comparable and conclusions drawn from analysis of group means may not be accurate.

Overall, the findings of the present study support the importance of obtaining and analyzing mothers' and fathers' reports about their children separately to assess treatment outcome. As with pretreatment or diagnostic measures, relying on only one parent's report or averaging parents' reports may obscure important information. Moreover, relying on an average parent report may lead to an overestimation of treatment effects when calculating effect size.

The findings of the present study suggest that the relationship between changes in parents' ratings and therapists' ratings of improvement may be a function of the type of treatment employed and which parent's report is considered. In the present study, correlations between parents' proportional gain scores and therapists' ratings of improvement were only significant for fathers' and average parent ratings, and only if parents were involved in treatment. These findings suggest the hypothesis that parents' ratings may change both as a func-

tion of changes in their child and changes in themselves as a result of treatment. Moreover, this may be the case only for fathers' ratings. Additionally, these data point to the need for further investigation into the factors affecting reliability and agreement between different sources of outcome data, in this case between parents and the therapist.

Multiple Regressions

Both parents' ratings were influenced by the type of problem manifested by their child. Staffing ratings of the degree of behavioral and emotional disorder displayed by the child accounted for significant amounts of variance in each parents' ratings of positive behaviors and in fathers' ratings of negative behaviors. Staffing ratings which showed a child's disorder to be more emotional than behavioral were associated with higher ratings for each parent on positive items and lower ratings for fathers on negative items. The direction of this relationship suggests that parents may be more able to see positive behaviors in children with emotional rather than behavioral problems. Additionally, fathers may rate emotional problems less severely on negative items. This may be related to behavioral disorders being more overtly difficult for parents than emotional disorders, which may be more covertly manifested. That is, behavioral disorders, which are related to more overt behavior disturbance may be more difficult for parents and result in less recognition of the child's positive behaviors and a more severe rating on negative behaviors. This hypothesis should be examined in relation to individual parental differences and actual differences between emotional and behavioral disorders.

The findings of the present study suggest that mothers' and fathers' perceptions of their children are influenced by somewhat different factors. This is consistent with the findings of Ferguson at al. (1974) that mothers and fathers are sensitive to different behavioral dimensions. This lends further support to the contention that mothers' and fathers' reports should be considered separately. In general, it appears that fathers' perceptions may be more influenced by factors related to their own response to their child's problems than are mothers' perceptions. For example, being embarassed by their child's problems accounted for a significant amount of the variance in fathers', but not mothers' ratings.

The variables accounting for significant amounts of variance in D scores appear to be more related to parent factors than child factors. The only child factor entering the regression equation for D scores was Age. It would appear that differences between parents' perceptions of their children may be related more to themselves than to their child's behavior. For example, if mothers reported that they felt their child had an emotional problem their ratings were higher than fathers'. Similarly, if fathers reported that they felt their child had an emotional problem their ratings were higher than mothers'. It should be noted that this finding is based on overall group means. For this to be more meaningful a more detailed analysis of individual parent dyads should be undertaken.

Three variables accounted for significant amounts of variance of interparent agreement as measured by the AD score: (a) the quality of the marital relationship, as measured by staffing ratings of the need for marital treatment; (b) whether the presenting problem was overt or covert, as rated by the author and one other rater; and (c) whether the child had been previously evaluated, as reported by the parents. The findings for the first 2 of these variables are consistent with the results of other studies suggesting that these are important factors affecting parents' perceptions and interparent agreement. Previous studies have suggested that the quality of the marital relationship may affect parents perceptions of their children (e.g., Christensen et al., 1983). The present measure of the quality of the marital relationship, staff ratings of parents' need for marital therapy, did not account for a significant amount of variance in either parents' individual perception but accounted for 15% of the variance in interparent agreement. This suggests that although the quality of the marital relationship may affect parents' individual perceptions, it has a more direct effect on interparent agreement. This would suggest that it may be most important to consider parents' reports separately when there is a high degree of marital discord. Different studies, using different measures of the quality of the marital relationship, have led to different conclusions regarding its effect on parents' perceptions of their children, however. The present findings regarding the impact of the quality of the marital relationship on interparent agreement need to be cross validated using

different measures of marital discord, including one from the parents, and different samples. Similarly, the effect of the presenting problem rating along the overt/covert dimension is consistent with the findings of other studies (e.g., Quay, 1977). Interparent agreement appears to be generally better for more overt behaviors. However, most of the distinctions between overt and covert behaviors have been based on checklists with separate subscales for measuring internalizing versus externalizing behaviors separately. The Behavior Problem Checklist, for example, has separate scales for Personality Problems (PP) and Conduct Problems (CP). The CP scale, which measures more overt behaviors, has been shown to have consistently higher interparent correlations (Quay, 1977) and to be related to higher levels of interparent agreement (Jacob et al., 1982).

It is also of interest to note the variables that did not affect interparent agreement, although some of them were related to the mothers' and fathers' individual perceptions as discussed above. Of the child factors, the child's age and sex, the duration and severity of his or her disorder, and whether the problem occured at home or only at school did not account for a significant amount of variance in interparent agreement. Of the parent factors investigated, parents' need for, or mothers' history of, individual treatment; the quality of the parents' child rearing techniques; the parents' level of motivation; parents' level of education; and several measures of parents' cognitions and expectations about their child's problems did not affect interparent

agreement. Similarly, referral source did not account for a significant amount of variance in interparent agreement.

Internal Validity

Several limitations of the present study should be discussed. Three issues relate to the WSCL itself. First, the WSCL has not been a highly researched instrument, and the demonstrations of its reliability and validity have been limited to investigations by the developer of the instrument (Wimberger & Gregory, 1968; Wimberger & Millar, 1968). The present study does lend support to Wimberger and Gregory's (1968) findings regarding acceptable levels of reliability for the WSCL. Secondly, the WSCL does not have subscale or factor scores that would enable more specific analyses based on contemporary diagnostic distinctions between internalizing and externalizing disorders. Much of the research regarding interparent agreement has utilized checklists with separate factor scores which have been shown to affect interparent agreeement and reliability (e.g., Quay, 1977). The overt/covert dimension used to rate presenting problems in the present study was based on judgements made by the author and one other rater based upon a description of the presenting problem(s). These judgements, although based on distinctions made in the literature, were often difficult to make. Future judgements should be based systematically on more rigorous criteria such as the use of checklists with separate scales. Lastly, only total WSCL scores were analyzed. Ideally, an item analysis would have been utilized to assess interparent agreement on specific items. Unfortunately, only total scores were available in the archival data used in the present study.

Several other limitations should be mentioned. The present study made a distinction between positive and negative item scores in an attempt to make the total WSCL scores more specific and interpretable. This distinction has not been regularly made in the literature and its use, although apparently valuable requires cross-validation. Some studies have mentioned differences between mothers and fathers regarding their ratings of more positive behaviors such as socialization (e.g., Thompson & McAdoo, 1973). Because mothers and fathers appear to be sensitive to different behavioral dimensions, distinctions along a positive/negative dimension may be informative. Additionally, the questions at the end of the WSCL used to assess parents' cognitions and expectations about their child's disorder cannot be interpreted as variables independent from parents WSCL scores. These variables were occasionally significant for individual parents' perceptions but not for interparent agreement. The use of independent measures of parental cognitive factors should be undertaken to examine further the potential relationship between interparent agreement and mother's and father's cognitions and expectations regarding their child's behavior.

Finally, although some of the variables used in the present study accounted for statistically significant amounts of variance, these amounts were often minimal. This raises the question of the clinical significance of some of the present findings. The amount of variance in mother's and fathers' criterion variables ranged from 8%, for mothers' negative ratings, to 34% for fathers' negative ratings. Similarly, the

amount of variance accounted for in D and AD scores ranged from 7% to 26%. Even the highest of these percentages, 34%, leaves a great deal of unexplained variance. This suggests either that factors other than those investigated by the present study account for the remainder of the variance or that the operational definitions and measures of the variables used were inadequate. In order to rule out the latter possibility it would be necessary to examine the internal reliability of the current measures and to validate the current measures against data obtained from other measures and sources. For example, staffing ratings on the nature of the child's disturbance could be compared to parent ratings on checklists with separate scale scores or to descriptions of the child's disorder from other sources. Similarly, staffing ratings of the need for marital treatment could be compared with parents' self report on a marital satisfaction inventory. It is particularly noteworthy that parents' reponses to the question "Does your spouse agree that there are problems?" in regard to their child's behavior did not account for more variance in AD scores, a measure of agreement. This raises some question about the relationship between parents' perceptions about agreement with their spouse and statistical definitions of interparent agreement calculated from differences in their checklist scores.

Areas for Future Research

One important area for future research is the analysis of the contingencies of interparent agreement. That is, the factors affecting mothers' and fathers' ratings of their children need to be investigated

in terms of one parent's response as a function of the reponse of the other parent. The one variable investigated in this manner in the present study (the question: "Does your spouse agree that there are problems?") did account for a significant amount of variance in interparent agreement. The present study included parents' responses to several other questions related to their cognitions and expectations about their child's behavior. However, these responses were analyzed independently for mothers and fathers. Parents' responses to several of these questions accounted for significant amounts of variance in the ratings on the WSCL. Whether or not parents' think their child has an emotional problem or are embarassed by it appears to influence their ratings of the child. An analysis that accounted for contingencies between parents' responses to such questions would provide important additional information about interparent agreement on such ratings. For example, if a father is embarassed by his child's problem does it make any difference if the mother is or isn't embarassed. That is, what are the effects of a father's embarassment about his child's problem on his rating of the child and upon interparent agreement contingent upon his spouse's embarassment or lack thereof. Another area needing additional research is the degree of reliability and agreement between parents' reports and information obtained from different sources (e.g., therapists, teachers, and the child him or herself). The present study suggests that the type of treatment approach employed may effect the degree of correlation between parents' and therapists' reports. Elucidation of

factors affecting this relationship will aid in the assessment of treatment outcome for child psychotherapy.

Summary

Conclusions about interparent agreement based on differences between group means may be misleading due to high levels of variability for individual parental dyads. The findings of the present study suggest that different factors affect mothers' and fathers' perceptions and levels of interparent agreement. Specifically, the present study supported the contention that levels of interparent agreement are affected by the qualtity of the marital relationship and the type of disorder manifested by the child. Additional research using measures of agreement such as the AD score recommended by Jacob et al., (1982) and utilized in the present study is still needed to confirm what factors affect levels of interparent agreement. Moreover, interactions between relevant factors such the child's age and sex, and the duration and severity if the child's disorder should be explored. Such information will be beneficial to researchers and clinicians alike in the determination of the costs and benefits of using one or both parents as a source of information about their children. At present, the most prudent conclusion appears to be to obtain and analyze parents' reports separately whenever feasable.

REFERENCES

Abramovitz, C.V. (1976). The effectiveness of group psychotherapy with children. Archives of General Psychiatry, 33, 320-326.

Achenbach, T.M., & Edelbrock, C.S. (1978). The classification of child psychopathology: A review and analysis of empirical efforts. Psychological Bulletin, 85, 1275-1301.

Achenbach, T. M., & Edelbrock, C. S. (1979). The child behavior profile II: Boys aged 12-16, and girls aged 6-11 and 12-16. Journal of Consulting and Clinical Psychology, 47, 223-233.

American Psychiatric Association, Commission on Psychotherapies. (1982). Psychotherapy research: Methodological and efficacy issues. Washington, DC: Author.

Atkeson, B.M., & Forehand, R. (1978). Parent behavioral training for problem children: An examination of studies using multiple outcome measures. Journal of Abnormal Child Psychology, 6, 449-460.

Barrett, C.L., Hampe, I.E., & Miller, L. C. (1978). Research on child psychotherapy. In S. L. Garfield & A. E. Bergin (Eds.), Handbook of psychotherapy and behavior change: An empirical analysis (2nd. Ed.). New York: John Wiley and Sons.

Burrows, K, Rogers-, & Kelley, C. K. (1983). Parental interrater reliability as a function of situational specificity and familiarity of target child. Journal of Abnormal Child Psychology, 11, 41-48.

Cass, L. K., & Thomas, C. B. (1979). Childhood pathology and later adjustment: The question of prediction (chap. 3, Review of outcome research with children). New York: John Wiley and Sons.

Christensen, A. Phillips, S., Glasgow, R. E., & Johnson, S. M. (1983). Parental characteristics and interactional dysfunction in families with child behavior problems: A preliminary investigation. Journal of Abnormal Child Psychology, 11, 153-166.

Clement, P. W., Fazzone, R. A., & Goldstein, B. (1970). Tangible reinforcers and child group psychotherapy. American Academy of Child Psychiatry, 9, 409-427.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences (rev. ed.). New York: Academic Press.

Colletti, G., & Harris, S. L. (1977). Behavior modification in the home: Siblings as behavior modifiers, parents as observers. Journal of Abnormal Child Psychology, 5, 21-30.

Cronbach, L. J., & Furby, L. (1970). How should we measure "change"- or should we? Psychological Bulletin, 74, 68-80.

Davidson, C. V., & Davidson, R. H. (1983). The significant other as data source and data problem in psychotherapy research. In M. J. Lambert, E. R. Christensen, & S. S. DeJulio (Eds.), The assessment of psychotherapy outcome. New York: John Wiley and Sons.

Dubey, D.R., O'Leary, S. G., & Kaufman, K. F. (1983). Training parents of hyperactive children in child management: A comparative outcome study. Journal of Abnormal Child Psychology, 11, 229-246.

Edelbrock, C. (1983). Problems and issues in using rating scales to assess child personality and psychopathology. School Psychology Review, 12, 293-299.

Edelbrock, C. S., & Achenbach, T. M. (1980). A typology of Child Behavior Profile patterns: Distribution and correlates for disturbed children. Journal of Abnormal Child Psychology, 8, 441-470.

Emery, R. E., & O'Leary, K. D. (1984). Marital discord and child behavior problems in a non-clinic sample. Journal of Abnormal Child Psychology, 12, 411-420.

Eyberg, S. M., & Johnson, S. M. (1974). Multiple assessment of behavior modification with families: Effects of contingency contracting and order of treatment. Journal of Consulting and Clinical Psychology, 42, 594-606.

Ferguson, L.R., Partyka, L.B., & Lester, B. M. (1974). Patterns of child perception differentiating clinic from non-clinic children. Journal of Abnormal Child Psychology, 2, 169-181.

Fine, S., Knight-Webb, G., & Breau, K. (1976). Volunteer adolescent group therapy: Effects on patients and volunteers. British Journal of Psychiatry, 129, 407-413.

Firestone, P., Kelly, M. J., & Fike, S. (1980). Are fathers necessary in parent training groups? Journal of Clinical Child Psychology, 9, 44-47.

Fiske, D. W. (1975). The use of significant others in assessing the outcome of psychotherapy. In I. E. Waskow & M. B. Parloff (Eds.), Psychotherapy change measures. Washington, DC: Department of Health Education and Welfare.

Forehand, R., Griest, D. L., & Wells, K. C. (1979). Parent behavioral training: An analysis of the relationship among multiple outcome measures. Journal of Abnormal Child Psychology, 7, 229-242.

Forehand, R., Wells, K. C., McMahon, R. J., Griest, D., & Rogers, T. (1982). Maternal perception of maladjustment in clinic referred children: An extension of earlier research. Journal of Behavioral Assessment, 4, 145-151.

Goldman, J., L'Engle-Stein, C., & Guerry, S. (1983). Psychological methods of child assessment. New York: Brunner Mazel, Inc.

Graham, P., & Rutter, M. (1968). The reliability and validity of the psychiatric assessment of the child. II Interview with the parent. British Journal of Psychiatry, 114, 581-592.

Griest, D. L. & Wells, K. C. (1983). Behavioral family therapy with conduct disorders in children. Behavior Therapy, 14, 37-53.

Guerney, B.G., Shapiro, G. B., & Stover, L. (1968). Parental perceptions of maladjusted children: Agreement between parents and relation to mother-child interaction. Journal of Genetic Psychology, 113, 215-225.

Herjanic, B., & Reich, W. (1982). Development of a structured psychiatric interview for children: Agreement between children and parents on individual symptoms. Journal of Abnormal Child Psychology, 10, 307-324.

Horton, L. (1984). The father's role in Behavioral Parent Training: A review. Journal of Clinical Child Psychology, 13, 274-279.

Humphreys, L. E., & Ciminero, A. R. (1979). Parent report measures of child behavior: A review. Journal of Clinical Child Psychology, 8, 56-63.

Irwin, E., Levy, P., & Shapiro, P. (1972). Assessment of drama therapy in a child guidance setting. Group Psychotherapy and Psychodrama, 25, 105-116.

Jacob, T., Grounds, L., & Haley, R. (1982). Correspondence between parents' reports on the Behavior Problem Checklist. Journal of Abnormal Child Psychology, 10, 593-608.

Jesness, C. F. (1975). Comparative effectiveness of Behavior Modification and Transactional Analysis programs for delinquents. Journal of Consulting and Clinical Psychology, 43, 758-759.

Johnson, S. M., & Christensen, A. (1975). Multiple criteria follow-up of behavior modification with families. Journal of Abnormal Child Psychology, 3, 135-154.

Johnson, S. M., & Eyberg, S. (1975). Evaluating outcome data: A reply to Gordon. Journal of Consulting and Clinical Psychology, 43, 917-919.

Kazdin, A. E., & Heidish, I. E. (1984). Convergence of clinically derived diagnoses, and parent checklists among inpatient children. Journal of Abnormal Child Psychology, 12, 421-436.

Kendall, P. C., & Finch, A. J., Jr. (1978). A Cognitive-Behavioral treatment for impulsivity: A group comparison study. Journal of Consulting and Clinical Psychology, 46, 110-118.

Kissel, S. (1974). Mothers and therapists evaluate long term and short term child therapy. Journal of Clinical Psychology, 30, 296-299.

Kogan, K. K., & Gordon, B. N. (1975). A mother instruction program: Documenting change in mother-child interaction. Child Psychiatry and Human Development, 5, 189-200.

Lessing, E. E., Black, M., Barbera, L., & Seibert, F. (1976). Dimensions of adolescent psychopathology and their prognostic significance for treatment outcome. Genetic Psychology Monographs, 93, 155-168.

Lessing, E. E., & Clarke, C. C. (1982). Reliability and validity of IJR Behavior Checklistscores: Number versus pathology level of symptoms. Journal of Abnormal Child Psychology, 10, 337-362.

Leventhal, T., & Weinberger, G. (1975). Evaluation of a large scale brief therapy program for children. American Journal of Orthopsychiatry, 45, 119-133.

Lindholm, B. W., & Touliatos, J. (1982). Checklist agreement among observers of children. Psychology in the Schools, 19, 548-551.

Lundeen, R. (1977). Modifying children's misbehavior through parent workshops. Journal of Clinical Child Psychology, 6, 36-37.

Marvit, R., Lind, J., & McLaughlin, D. (1974). Use of video-tape to induce attitude change in delinquent adolescents. American Journal of Psychiatry, 131, 996-999.

Miller, L. C. (1964). Q-sort agreement among observers of children. American Journal of Orthopsychiatry, 34, 71-75.

Mintz, J., Luborsky, L. & Christoph, P. (1979). Measuring the outcome of psychotherapy: Findings of the Penn Psychotherapy Project. Journal of Consulting and Clinical Psychology, 47, 319-334.

Novick, J. (1965). Comparison between short-term group and individual therapy in effecting change in nondesirable behavior in children. International Journal of Group Therapy, 15, 366-373.

Novick, J., Rosenfeld, E., & Bloch, D. A. (1966). Situational variation in the behavior of children. Journal of Consulting Psychology, 30, 488-493.

Novick, J., Rosenfeld, E., & Bloch, D. A., & Dawson, D. (1966). Ascertaining deviant behavior in children. Journal of Consulting Psychology, 30, 230-238.

O'Leary, K. D., & Turkewitz, H. (1978). Methodological errors in marital and child treatment research. Journal of Consulting and Clinical Psychology, 46, 747-758.

O'Leary, K. D., Turkewitz, H., & Taffel, S. J. (1973). Parent and therapist evaluation of behavior therapy in a child psychological clinic. Journal of Consulting and Clinical Psychology, 41, 279-283.

Patterson, G. R., Cobb, J. A., & Ray, R. S. (1973). A social engineering technology for retraining the families of aggressive boys. In H. E. Adams & I. P. Unikel (Eds.), Issues and trends in Behavior Therapy. Springfield, IL: C. C. Thomas.

Peterson, D. R. (1961). Behavior problems in middle childhood. Journal of Consulting Psychology, 25, 205-209.

Peterson, D. R., Becker, W. C., Shoemaker, D. J., Luria, Z., Hellmer, L. A. (1961). Child behavior problems and parental attitudes. Child Development, 32, 151-162.

Porter, B., O'Leary, K. D. (1980). Marital problems and childhood behavior problems. Journal of Abnormal Child Psychology, 8, 287-295.

Quay, H. C. (1977). Measuring dimensions of deviant behavior: The Behavior Problem Checklist. Journal of Abnormal Child Psychology, 5, 277-287.

Rickard, K. M., Graziano, W. & Forehand, R. (1984). Parental expectations and childhood deviance in clinic referred and non-clinic referred children. Journal of Clinical Child Psychology, 13, 179-186.

Robinson, E. A., & Anderson, L. L. (1983). Family adjustment, parental attitudes and social desirability. Journal of Abnormal Child Psychology, 11, 247-256.

Ross, A. W. Behavior therapy with children. (1978). In S. L. Garfield & A. E. Bergin (Eds.) Handbook of Psychotherapy and behavior change (2nd. Ed.). New York: John Wiley and Sons.

Sines, J. O., Paulker, J. D., Sines, L. K., & Owen, D. R. (1969). Identification of clinically relevent dimensions of children's behavior. Journal of Consulting and Clinical Psychology, 33, 728-734.

Speer, D. C. (1971). Behavior Problem Checklist (Peterson-Quay): Baseline data from parents of child guidance and non-clinic children. Journal of Consulting and Clinical Psychology, 36, 221-228.

Strupp, H. H., & Hadley, S. W. (1977). A tripartite model of mental health and therapeutic outcomes: With special reference to negative outcomes in psychotherapy. American Psychologist, 32, 187-196.

Szapocznik, J. Kurtines, W. M., Foote, F. H., Perez-Vidal, A., & Hervis, O. (1983). Conjoint versus one-person family therapy: Some evidence for the effectiveness of conducting family therapy through one person. Journal of Consulting and Clinical Psychology, 51, 889-899.

Taylor, W. F., & Hoedt, K. C. (1974). Classroom-related behavior problems: Counselors, parents, teachers and children. Journal of Counseling Psychology, 21, 3-8.

Thompson, R. J., & Bernal, M. E. (1982). Factors associated with parent labelling of children referred for conduct problems. Journal of Abnormal Child Psychology, 10, 191-202.

Thompson, R. J., & Curry, J. F. (1983). A construct validity study of the Missouri Childrens Behavior Checklist with developmentally disabled children. Journal of Clinical Psychology, 39, 691-695.

Thompson, R. J., & McAdoo, W. G. (1973). A comparison of mother's and father's checklist ratings of outpatient boys and girls. Journal of Community Psychology, 1, 387-389.

Tinsley, H. E. A., & Weiss, D. J. (1975). Interrater reliability and agreement of subjective judgements. Journal of Counseling Psychology, 22, 358-370.

Tramontana, M. G., & Sherrets, S. D. (1983). Assessing outcome in disorders of childhood and adolesence. In M. J. Lambert, E. R. Christensen, & S. S. DeJulio (Eds.), The assessment of psychotherapy outcome. New York: John Wiley and Sons.

Tuma, J. M., & Sobotka, K. R. (1983) Traditional therapies with children. In T. H. Ollendick & M. Hersen (Eds.), Handbook of child psychopathology. New York: Plenum.

Vandenbos, G. R., & Pino, C. D. (1980). Research on the outcome of psychotherapy. In G. R. Vandenbos (Ed.), Psychotherapy: Practice, research, policy. Beverly Hills, CA: Sage.

Wimberger, H. C., & Gregory, R. J. (1968). A behavior checklist for use in child psychiatry clinics. Journal of the American Academy of Child Psychiatry, 7, 677-688.

Wimberger, H. C., & Millar, G. (1968). The psychotherapeutic effects of initial clinic contact on child psychiatry patients. In S. E. Lesse (Ed.), An evaluation of the psychothrapies, Springfield, IL: Thomas.

Wolf, T. M. (1981). Measures of deviant behavior, activity level and self concept for educable mentally retarded-emotionally disturbed students. Psychological Reports, 48, 903-910.

Zold, A. C., & Speer, D. C. (1971). Follow up study of child guidance patients by means of the Behavior Problem Checklist. Journal of Clinical Psychology, 27, 519-524.

APPENDIX A

STAFF RATING FORM

CASE EVALUATIONS

Your Name: _____ Date: _____ Referred Child: _____

Status of Case: Pre-treatment In-treatment (time) _____ Post-treatment

Contact with case:

THERAPY SOCIAL HISTORY DIAGNOSTIC STAFFING SUPERVISION

OTHER (specify) _____ (circle each item that applies)

Make ratings along an 0 to 6 scale where 0 - 'not at all' and 6 = 'very much' or 'very great'.

RATINGS

_____ PRESENCE OF MBD SYMPTOMS. Rate according to the presence of the six cardinal features of MBD. (1) Motor behavior (hyperactive); (2) attentional and perceptual-visual problems' (3) learning difficulties; (4) impulse control; (5) problems in interpersonal relations; (6) emotionality (labile and overreactive). Rate one point for the appearance of each symptom.

_____ DEGREE OF BEHAVIORAL MANAGEMENT PROBLEM: Rate according to the quantity of specific behaviors which need to be changed. e.g. social disruptions in school, poor academic performance, wets bed, refuses to listen to parents' commands, lies, steals, fights, etc.

_____ DEGREE OF EMOTIONAL PROBLEM: Rate according to presence of internal emotional state (not directly observable) such as poor self-concept, anxiety, depression, identity problems, and/or degree of inadequate or disorganized personality development.

_____ INEFFECTIVENESS OF PARENTS CHILD REARING TECHNIQUES. Rate according to inability of parents to set limits and enforce them, and to discipline and reward child approrpiately.

_____ NEED FOR MARITAL/INDIVIDUAL THERAPY. Rate according to degree to which parents need some kind of therapy/counseling for themselves not directly related to referred child's probelm(s).

_____ LACK OF MOTIVATION TO SEEK HELP. (Higher scores reflect less motivation).

APPENDIX B

	NEVER	SELDOM	FRE-QUENTLY	VERY OFTEN
35. Talks back to parents				
36. *Has been held back a grade in school				
37. Lacks self-confidence				
38. *Has been in trouble with Juvenile Authorities				
39. Has sleeping disturbances				
40. Prefers to play with children not his/her age				
41. Cries easily				
42. Refuses parental instructions				
43. Gets along poorly with children of opposite sex				
44. Is irritable				
45. Gets along well with grownups				
46. Has speech difficulty				
47. Gets along poorly with brothers and sisters				
48. Is resentful of discipline				
49. Teases others				
50. Is fearful				
51. Is stubborn				
52. Is nervous and jumpy				
53. Is bossy				
54. Is destructive				
55. Is overactive				
56. Is afraid to defend herself/himself				
57. Has physical complaints				
58. Wets bed				
59. Sucks thumb				
60. Bites nails				
61. Masturbates				
62. Shows unusual interest in fires				
63. Has a tic (nervous twitch)				
64. Does not show feelings				
65. Is concerned about neatness				
66. Complains about going to school				
67. Other problems not listed:				

The next nine questions are directed to you, as the child's parents. They may not be exactly appropriate to your special situation, but please answer them to the best of your ability.

	YES	UNDECIDED	NO
68. Do you think that your child has an emotional problem?			
69. Does it embarass you that your child has an emotional problem?			
70. Does your wife/husband agree that there are problems?			
71. Do you feel in part responsible for your child's problems?			
72. Do you feel that your child will outgrow the problem?			

APPENDIX C

Relevant Correlations Between Mothers' and Fathers' Pretreatment Checklist Scores and Predictor Variables

	Criterion Variables			
	Mother		Father	
Predictor Variables	Positive Items	Negative Items	Positive Items	Negative Items
Child				
Age	-.16			.28**
Sex				
Duration			-.18*	
Severity		.29**	-.21*	.27**
Emot/Beh Prob	-.20*		-.23*	.28**
Overt/Covert				
Home/School				
Prev. Eval.				.24**
Mother				
Rx Need				
Prev. Rx				
CRT		.27**		
Motivation				
High School	.32**	-.23*		
College				
Question A				
Question B				
Question C				
Question D				
Question E				
Father				
Rx Need				
CRT				.19*
Motivation				
High School			.20*	
College				
Question A				.38**
Question B			-.18	.22**
Question C			-.26**	
Question D				
Question E				-.29**
Other				
Marital Rx		.19*		
Spouse Agree?				
Referral				

Note. Questions: A "Do you think your child has an emotional problem?"; B "Does it embarass you?"; C "Does your spouse agree?"; D "Do you feel in part responsible?"; E "Do you think your child will outgrow the problem?". Abbreviations: (CRT) Child Rearing Techniques.
*p<.05. **p<.01.

APPENDIX D

Relevant Correlations Between Difference and Absolute Difference Scores and Predictor Variables

Predictor Variables	Criterion Variables: Difference: Positive Items	Difference: Negative Items	Absolute Difference: Positive Items	Absolute Difference: Negative Items
Child				
Age	-.21*			
Sex				
Duration		-.19*		
Severity				
Emot/Beh Prob				.39**
Overt/Covert				-.24*
Home/School				
Prev. Eval.				.15
Mother				
Need Rx				
Prev. Rx				.27**
CRT				
Motivation				
High School	.28**			
College		-.20*		
Question A		.14		.22*
Question B				
Question C				
Question D				
Question E				
Father				
Need Rx				
CRT				
Motivation				
High School				
College		-.23*		
Question A		-.16		
Question B	.20**	-.30**		.22*
Question C				
Question D				
Question E				
Other				
Marital Rx				.39**
Spouse Agree?			.27*	
Referral		.19*		

Note. Questions: A "Do you think your child has an emotional problem?"; B "Does it embarass you?"; C "Does your spouse agree?"; D "Do you feel in part responsible?"; E "Do you think your child will outgrow the problem?". Abbreviations: (CRT) Child Rearing Techniques.
*p<.05. **p<.01.

APPROVAL SHEET

The thesis submitted by Barry R. Lindstrom has been read and approved by the following committee:

Joseph A. Durlak, Ph.D., Director
Professor, Psychology
Loyola University of Chicago

Dan P. McAdams, Ph.D.
Associate Professor, Psychology
Loyola University of Chicago

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the thesis is now given final approval by the Committee with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Arts.

April 18, 1985
Date

Joseph A. Durlak
Director's Signature